

MARIA CECILIA SCHEFFER CONTIN

A VALIDADE DO TESTE N73 - TESTE DA PROFICIÊNCIA ORAL DA LÍNGUA

Dissertação para obtenção do grau de Mestre, Área de Concentração: Língua Inglesa, do Curso de Pós-Graduação em Letras. Setor de Ciências Humanas, Letras e Artes da Universidade Federal do Paraná.

CURITIBA
1981

PROFESSOR ORIENTADOR

Doutora Otília Arns

Titular de Língua e Literatura Inglesa
da Universidade Federal do Paraná.

AGRADECIMENTOS

Ao Prof. Michael Watkins, Professor Visitante de Lingüística Aplicada na UFPR, pelas suas valiosas sugestões.

Aos demais que, de uma forma ou outra, contribuíram para a realização deste trabalho.

SUMÁRIO

	PÁGINA
Resumo	vii
Abstract	ix
1. Introdução	1
1.1. Apresentação do Trabalho	1
1.2. Hipótese	2
1.3. Metodologia	2
1.4. Objetivos	3
2. Revisão da Literatura	4
2.1. O Ato da Comunicação Oral	4
2.1.1. O modelo de comunicação oral apresentado por LEESON	7
2.2. Testes e seus Conceitos	19
2.2.1. Propósito, forma e conveniência de tes- tes	19
2.2.2. Critérios para comparação com relação à eficiência - confiabilidade, validade e correlação	22
2.2.3. Contagem de pontos e aplicação	26

2.2.4. Estágios pelos quais passaram os testes relacionados à língua nos últimos anos - Relação entre testes, objetivos e currículos	27
2.2.5. Testes integrativos & testes analíticos - Participantes & não-participantes	31
2.2.6. Proficiência oral	34
2.2.7. Diferentes maneiras de medir a proficiência oral	35
2.3. <i>Test N73: Instantaneous Evaluation of Speaking Ability</i>	54
2.4. Aplicações dos Conceitos ao Teste N73 ..	62
Notas	66
3. O Experimento	68
Nota	74
4. Análise Estatística dos Escores - Comparação dos Testes	75
4.1. Resultado do Falante Nativo	75
4.2. Resultado do Teste N73	77
4.2.1. Variações de Fluência e Exatidão - Transformação de Escalas	78
4.2.2. Notas obtidas com N73	81
4.3. Apresentação dos Dados	85
4.4. Aplicação do Coeficiente de Correlação por Postos de Kendall - T (TAU)	87
4.4.1. Cálculo de S	89
4.4.2. Cálculo de Tx e Ty	91

	PÁGINA
4.4.3. Cálculo de T e Prova de Significância (z)	92
5. Conclusão	95
6. Referências Bibliográficas	97

RESUMO

O presente trabalho desenvolve uma comparação entre dois métodos de avaliação da proficiência oral de um indivíduo. Os dois métodos utilizados foram a entrevista oral realizada por um falante nativo e um novo método proposto por Nicolas Ferguson - *Test N73: Instantaneous Evaluation of Speaking Ability*.

Esta comparação foi realizada com o intuito de verificar se o teste N73 é tão confiável quanto a entrevista oral realizada por falantes nativos. Para tanto, testamos 40 alunos de inglês matriculados na Sociedade Brasileira de Cultura Inglesa em ambos os métodos e gravamos as amostras para uma melhor avaliação.

Uma vez que para a concretização do experimento fazia-se necessário termos um segundo teste, o qual iria servir de base para uma comparação com o N73, fizemos um levantamento das diferentes maneiras existentes de avaliar a habilidade oral, e dos conceitos relacionados com este tipo de teste. Desta mesma forma, visto que só estávamos interessada em testes de comu-

nicação oral, compusemos um relato sobre o que consiste o ato de fala.

Após a aplicação de cada teste individualmente, comparamos ambos os resultados para uma verificação do nível de correlação entre os dois métodos de avaliação da proficiência oral. Utilizando métodos estatísticos para a correlação, chegamos à conclusão de que realmente existe um relacionamento entre ambas as maneiras de testar, e fomos conduzida então, para a comprovação da hipótese do presente trabalho, ou seja, para a comprovação de que o teste N73 é tão confiável quanto a entrevista oral para testar a proficiência.

ABSTRACT

In this dissertation two methods of measuring oral proficiency are compared: one consisting of an interview with a native speaker; the other a new test put forward by Nicolas Ferguson - N73: *Instantaneous Evaluation of Speaking Ability*.

The purpose of this comparison was to establish whether N73 is as reliable as the oral interview with a native speaker. In order to do this we applied both tests to 40 students of the Anglo-Brazilian Cultural Society, making a tape recording of the samples.

Since a second type of test was needed for the purpose of the experiment, as a basis of comparison with N73, we examined the different tests of oral proficiency which exist, and the theories behind them. We also outlined the processes involved in the act of oral communication.

After applying each test individually, we compared both sets of results to discover the strength of the correlation between them, the conclusion being that

a positive correlation exists, and that the hypothesis, i.e. that the N73 test is as reliable as the interview for testing oral proficiency, has been proved.

1. INTRODUÇÃO

1.1. APRESENTAÇÃO DO TRABALHO

O presente trabalho procura fornecer uma avaliação de um novo teste, que se propõe medir a comunicação oral. Esta nova maneira de testar a comunicação oral foi elaborada por Nicolas Ferguson através do teste: *Test N73: Instantaneous Evaluation of Speaking Ability*.

Justificamos a escolha do tema através dos seguintes itens:

- a) O teste N73 pretende ser um método prático, rápido e objetivo para avaliar a comunicação oral.
- b) Visto que o N73 mede a fluência e exatidão da habilidade oral de um indivíduo em dois minutos, faz-se necessário testar sua confiabilidade e validade.
- c) De acordo com a opinião de vários metodologistas, pouco tem sido feito neste campo.

Após levarmos em consideração os fatos precedentes, decidimos verificar se o teste N73 corrobora a entrevista oral feita por falantes nativos, a qual, para o propósito do presente trabalho, é considerada confiável.

1.2. HIPÓTESE

O *Test N73: Instantaneous Evaluation of Speaking Ability* fornece uma avaliação da habilidade oral da língua, com resultados tão confiáveis quanto aqueles fornecidos através da entrevista realizada pelo falante nativo.

1.3. METODOLOGIA

Para a realização do experimento, tomamos como amostra 40 alunos matriculados no quinto ano da Sociedade Brasileira de Cultura Inglesa. Ambos os testes, o N73 e a avaliação pelo falante nativo foram aplicados aos mesmos alunos, sendo que para o N73 o tempo de avaliação foi de dois minutos e para a avaliação pelo falante nativo de cinco minutos.

1.4. OBJETIVOS

Para a execução deste trabalho, propusemo-nos alcançar os seguintes objetivos:

- a) descrever o que é a comunicação oral e quais os principais fatores com ela relacionados;
- b) examinar a bibliografia existente sobre a teoria e conceitos relativos a testes;
- c) relatar diferentes testes usados para medir a comunicação oral;
- d) descrever o teste N73;
- e) comparar o teste aplicado pelo falante nativo com o N73;
- f) analisar os resultados obtidos;
- g) concluir pela comprovação, ou não, da hipótese formulada.

2. REVISÃO DA LITERATURA

Como o teste a ser analisado se relaciona com a habilidade oral da língua, faremos, no presente capítulo, um relato sobre o ato da comunicação oral tomando como base LEESON¹. Na segunda parte da revisão da literatura, trataremos de diferentes tipos de testes existentes para medir a habilidade oral e critérios necessários para testar, para, então, apresentar o modelo do teste *N73: Instantaneous Evaluation of Speaking Ability*.

2.1. O ATO DA COMUNICAÇÃO ORAL

Para que um ato de fala complete seu percurso satisfatoriamente, muitos pré-requisitos são necessários, uma vez que diferentes fatores estão envolvidos no ato da comunicação oral.

Para que ocorra o ato de comunicação é necessário que pelo menos duas pessoas estejam envolvidas: o

emissor e o receptor. O emissor tem a necessidade interna de falar e ser ouvido. Portanto, cabe a ele o papel de codificar a mensagem, o de emití-la para alguém, que é o ouvinte, o receptor. Desta maneira, um teste deveria fazer com que o aluno sentisse esta mesma necessidade, isto é, sentisse necessidade de falar.

Ter algo para dizer e alguém para ouvir é o que OLLER chama de *pragmatic naturalness criteria*² (critério pragmático natural) da língua. O critério pragmático da língua é a situação normal do ato de comunicação, em que existe um falante que tem a necessidade de falar e ser ouvido.

OLLER faz a seguinte afirmação: *We need tests that meet the pragmatic naturalness criteria*² (p.305). Testes deveriam proporcionar aos alunos a chance de construírem os seus atos de fala, nos quais eles usassem as estruturas, nas devidas situações extralinguísticas. Devem apresentar situações que se assemelhem com o critério pragmático da comunicação.

Como a língua é primordialmente uma atividade oral, a comunicação oral é um elemento essencial da língua. Portanto, língua e comunicação deveriam ser ensinadas e testadas como um único elemento.

A comunicação ocorre, quando um elemento influencia outro elemento, isto é, o emissor influencia o receptor. Esta influência é conduzida através de um

canal, o qual por sua vez conduz a mensagem. Será de interesse para os requisitos do presente trabalho o canal oral: a comunicação levada da boca do emissor para o ouvido do receptor. Durante a comunicação, há a ocorrência de uma série de outros fatores extralingüísticos, que não deixam de ser relevantes para um ato bem sucedido. Estes fatores são: gestos, expressões faciais, manipulação com objetos e estados emocionais. Estados emocionais também influenciam a conversação cotidiana e podem afetar a participação de ambos: o emissor e/ou o receptor, nas suas funções de codificar ou decodificar a mensagem. Se tomarmos como exemplo a situação teste, verificaremos que uma grande parte dos examinandos terão a sua capacidade de comunicar-se, na língua em questão, mudada, devido ao fato de estarem sendo analisados, deixando, assim, o fator nervos entrar em ação.

Tanto o emissor como o receptor têm as suas respectivas funções na comunicação. Como já foi mencionado, a codificação, é a atividade do emissor, enquanto que o receptor tem a função de decodificar a mensagem. Portanto, quando alguém aprende uma segunda língua, ele tem que formar novos hábitos para codificar e decodificar, isto porque cada língua possui a sua maneira peculiar de comunicar as idéias.

2.1.1. O modelo de comunicação oral apresentado por LEESON

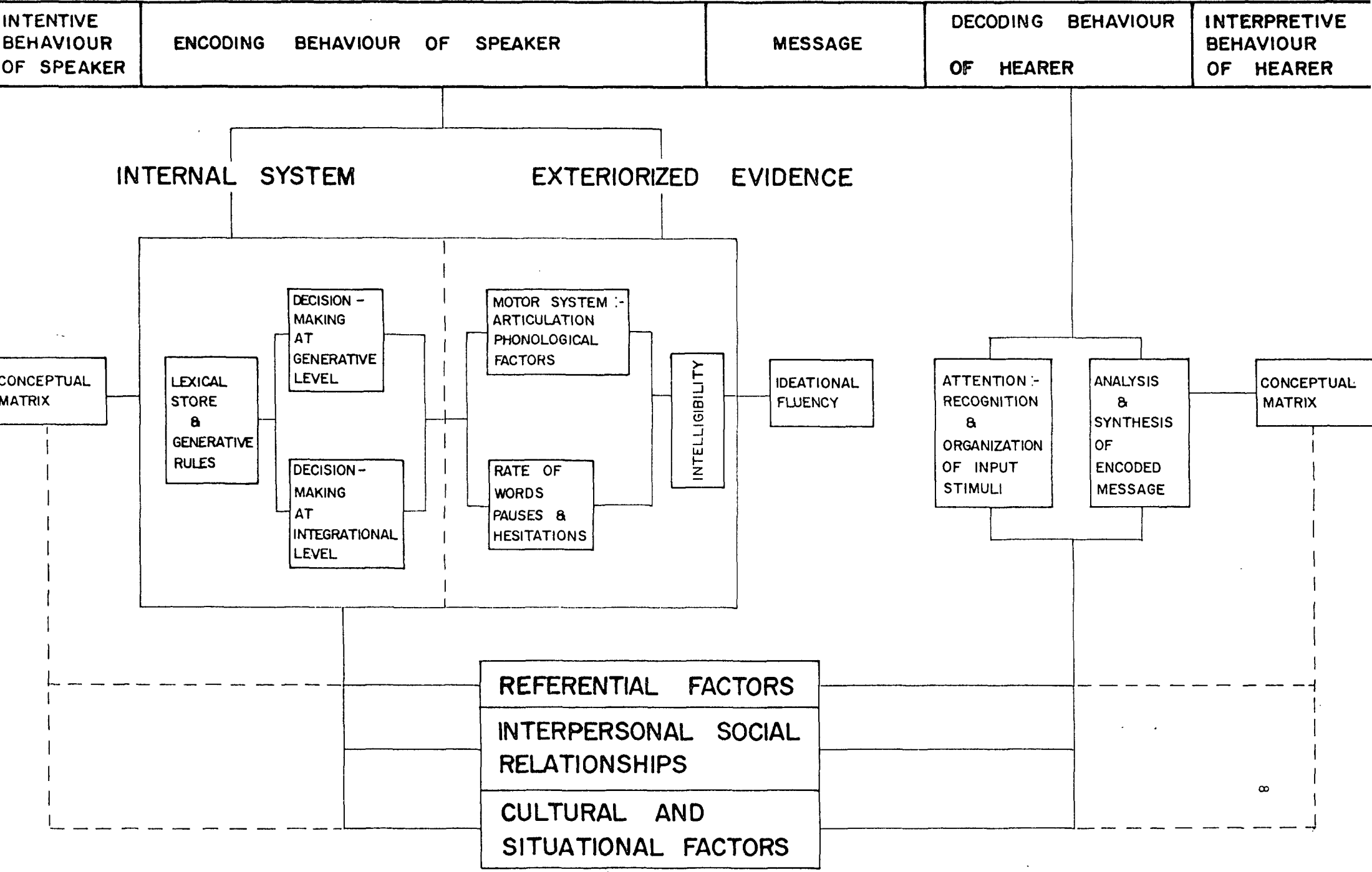
LEESON em sua obra *Fluency and Language Teaching* resume o que é a comunicação oral e o que acontece nela no seguinte quadro (p.30).

Baseando-nos no quadro de LEESON, e nas suas explicações verificamos que o ato de comunicação compreende cinco fases distintas:

- . comportamento intencional;
- . codificação;
- . mensagem;
- . decodificação;
- . interpretação.

O primeiro passo para a emissão da mensagem é o comportamento intencional do emissor, que tem a intenção de expressar alguma idéia, devendo, portanto, sentir-se motivado para este fim. Em seguida, vem a fase da formação das idéias, a codificação, denominada *encoding behaviour of speaker*.

A fase de codificação abrange um sistema interno e outro externo. Porém LEESON¹ afirma que *any separation of the encoding levels is an exercise of theoretical convenience and when one considers the immense*



complexity and speed of the articulatory motor responses involved at this level, the theoretical nature of such a division becomes even more clear (p.53).

A divisão da fase de codificação em sistema interno e externo apresentada por LEESON vem esclarecer quais os passos tomados na codificação de uma mensagem, o que ocorre com cada emissor ao emitir a sua idéia para os ouvintes. O sistema interno dos emissores contém o estoque léxico do indivíduo e as regras generativas que foram adquiridas por ele durante o aprendizado da língua. Com a intenção de comunicar-se, ele torna-se obrigado a optar por esta ou aquela regra generativa e por estes ou aqueles vocábulo, combinando ambos da melhor forma possível. Após a conclusão deste procedimento, o indivíduo passa para a exteriorização da mensagem, cuja fase apresenta os fatores fonológicos, pausas e hesitações, os quais deverão ser combinados entre si e com os elementos do sistema interno, inteligivelmente.

Para LEESON, saber uma língua, isto é, saber codificar em determinada língua é possuir a habilidade de responder a qualquer estímulo, treinado ou não treinado, de acordo com as regras fonológicas, sintáticas e semânticas pertencentes à língua em questão. A esta habilidade LEESON chama de *generative habit*¹ (p.6).

O hábito generativo consiste em apresentar ao

aluno um conjunto de regras, a partir das quais ele será capaz de formar um número infinito de sentenças. Contudo, para que ele possa gerar novas sentenças através do uso das regras generativas, é necessário um prévio e bom treinamento das mesmas. Partindo da presente premissa, ele declara que aprender uma língua é muito mais do que ser treinado como um papagaio, que só reproduzirá o que o fizeram repetir, tornando-se incapaz de manipular com a língua e adaptar-se a qualquer situação lingüística. Uma vez que as regras generativas podem ser aplicadas a um grande grupo de situações, LEESON propõe-nas como uma solução para estudantes estrangeiros, os quais sofrem as conseqüências da falta de tempo e falta de permanente contato com a língua.

Um dos problemas em torno da comunicação oral é a fluência. Para LEESON *fluency may be defined as the ability of the speaker to produce indefinitely many sentences conforming to the phonological, syntactical and semantic exigencies of a given natural language on the bases of a finite exposure to a finite corpus of that language*¹ (p.136).

Na opinião de LEESON, esta definição parece difícil de ser preenchida até mesmo por falantes nativos, visto que a qualidade do ato de comunicação dos falantes nativos apresenta um grande grau de variabilidade. A definição acima apresenta-se falha se consi-

derarmos que os falantes nativos que se caracterizam por apresentarem uma comunicação de baixa qualidade também são fluentes; eles têm a capacidade de formar sentenças, mas nem sempre de acordo com as regras estabelecidas pela estrutura da língua.

Relacionando a definição acima com o teste N73, verificamos que ela se divide em duas partes. A primeira parte da definição, ou seja a capacidade de produzir sentenças Nicolas Ferguson denomina *fluência*; enquanto que a segunda parte, ou seja, produzir sentenças de acordo com as regras determinadas pela língua ele chama de *exatidão*.

De início notamos que a fluência e conseqüentemente, a comunicação oral não são tão simples, podendo compreender muitos conceitos ao mesmo tempo.

A fluência compreende os três elementos da língua (fonológico, semântico e sintático), pausas e hesitações, que são elementos importantes e necessários para o ato de fala, e interação entre falante e ouvinte.

Com relação aos três elementos da língua, mencionamos diversos aspectos. Um elevado número de erros em qualquer um dos elementos é uma indicação de baixa fluência. No entanto, no que se refere ao aspecto semântico deve-se levar em consideração se a pessoa está falando sobre um assunto que domina, ou se está ten-

tando falar sobre tópicos desconhecidos. Falar sobre tópicos desconhecidos não revela o verdadeiro nível de fluência, mas sim a falta de conhecimento sobre o assunto, a falta talvez de formação intelectual.

Os elementos sintáticos e semânticos mantêm uma íntima relação no ato de fala. Forma e conteúdo são inseparáveis para a fluência, ao ponto de um depender do outro na formação do ato de fala. Para a combinação dos elementos semânticos, é necessário algum conhecimento da estrutura da língua, sendo que o contrário também é válido. O número de combinações possíveis no campo sintático não é tão grande comparado com as combinações do campo semântico. Quanto mais extenso for o componente semântico de uma pessoa maior será o seu número de possibilidades para transmitir uma idéia e, por sua vez, o componente semântico depende da formação intelectual dos indivíduos.

Partindo-se do princípio de que ambos os componentes, semântico e sintático, são de igual importância para o ato de fala e de que nada pode ser feito com somente um deles, é necessário que haja uma harmoniosa integração entre os dois, juntamente com o componente fonológico. Neste ponto, a proposição feita por HERRIOT³: *Language behaviour might consist of a hierarchy of skills rising from phonological to grammatical and semantic skills* (p.21) coincide com as idéias de

LEESON de que o ato de formulação é uma hierarquia de unidades.

Como importantes elementos do ato de fala, e agindo sobre a fluência, aparecem as pausas e hesitações. Enquanto que hesitações podem ser notadas facilmente, pausas não o são. Com relação à produção do ato de fala, LEESON afirmou que pausas são fatores relevantes para a fluência e causam grandes problemas para o aluno estrangeiro, uma vez que revelam a sua deficiência na fluência. A dificuldade de fazer pausas no devido lugar ocorre porque elas obedecem a uma regra mais ou menos flexível. Todos os falantes fazem pausas, porém cada falante apresenta a sua própria variação, que é um sinal de individualidade, de personalidade. Variações com respeito a pausas, isto é, fazer mais ou menos pausas pode ser inato ou pode ser resultante de uma habilidade adquirida através de treinamento.

Um ponto muito difícil ao comunicar-se em uma segunda língua é fazer pausas de acordo com os padrões estabelecidos pelos falantes nativos da língua, pois quando colocadas erroneamente elas são facilmente observadas.

Experimentos realizados com relação a pausas demonstraram que, para produzir enunciados superiores em qualidade, o falante deve fazer mais pausas do que

o normal.

Outros fatores extralingüísticos que influenciam tanto o processo de codificação como a fluência são os seguintes:

1. Fator referencial

O aspecto referencial consiste em aprender o significado de uma nova palavra através da referência desta palavra ao objeto; algumas vezes o aluno pode apresentar dificuldade em gravar certas palavras na língua dois, por ter dificuldade em associar tais palavras com o objeto. Isto pode acontecer devido ao fato de diferentes culturas apresentarem diferentes formas de objetos para descrever o mesmo conceito. Supondo-se que o objeto *mala* fosse diferente na cultura brasileira e na cultura inglesa, o aluno brasileiro na língua inglesa teria dificuldades em associar o conceito com o objeto real.

2. Embasamento cultural da comunidade

Culturas diferentes podem acarretar separação, e por conseguinte, a comunicação pode ser interrompida se um dos participantes não conhecer o embasamento cultural da comunidade. Existem hábitos e tradições que são comuns a todas as comunidades, mas, por outro lado, alguns são específicos de cada comunidade. É so-

bre estes hábitos específicos que ambos os participantes têm que estar cientes, no caso de pretenderem ter uma comunicação bem sucedida.

Outro aspecto com relação à cultura é a maneira como a comunidade fala; esta maneira varia de acordo com a região em que está estabelecida. Quando um indivíduo muda de uma região para outra de mesma língua, ele será capaz de comunicar suas idéias, porém seu acento, sua entonação, e até mesmo a pronúncia de algumas palavras será diferente, revelando que ele não pertence àquela região.

3. Diferentes funções sociais

Para diferentes situações, usam-se diferentes tipos de ato de fala. Certas vezes o falante tem que fazer certas considerações, tais como: para quem, aonde, e o que falar. De acordo com a pessoa com quem fala, o local no qual se encontra, ele escolherá uma certa maneira de comunicar a sua idéia, pois existem diferentes estilos para as diferentes classes sociais.

Quando alguém se dirige a um membro da família, esse alguém usa um tipo de linguagem diferente da que ele usa quando está no trabalho, assim como usará tipos diferentes para quando estiver em uma reunião social ou fazendo palestras.

A terceira fase no ato de comunicação é a men-

sagem. Ela é o elo que une o emissor ao receptor, transportando a idéia daquele.

A mensagem, segundo LEESON, é formada por fatores lingüísticos e extralingüísticos, e por pausas e hesitações. E ainda, todo ato de fala é composto por três elementos: o fonológico, o sintático e o semântico. O elemento semântico rege o que é para ser dito, enquanto os elementos sintáticos e fonológicos controlam como a mensagem é dita, mais especificamente, o sintático controla as estruturas e o fonológico os sons.

As duas últimas partes do modelo de LEESON são a decodificação e a interpretação da mensagem pelo receptor. Na decodificação o ouvinte reconhece, analisa e faz uma síntese da mensagem, decodificando-a antes de levá-la à fase da interpretação, em que ele vai entender o que foi emitido pelo falante, isto é, entender o objetivo da mensagem.

Embora o papel mais importante na comunicação seja o do emissor, o receptor é de grande importância para o emissor. Diremos que a interpretação do receptor chega algumas vezes a influenciar o falante, quando este está codificando, podendo ficar constrangido, ou não, em continuar o seu ato de fala. Constatamos, em HERRIOT³, a seguinte afirmação, a qual vem confirmar o que havíamos suposto: *The nature and behavior of the*

other part in the communication situation may also be a determinant of the nature of the language produced (p.151).

Pelas expressões do receptor nós podemos afirmar se ele está, ou não, entendendo a mensagem. No caso de resposta afirmativa, ainda nos restam algumas alternativas a serem indagadas: O que pensa ele a respeito do assunto? Ele aprova ou desaprova as idéias? As suas expressões faciais poderão encorajar o emissor a continuar ou poderão fazê-lo mudar de código. Elas demonstram se a mensagem atingiu o seu objetivo.

A comunicação bem sucedida implica a percepção e entendimento da mensagem. O emissor deve transmitir suas idéias da maneira mais clara possível, para que haja a devida reação por parte do receptor. Para cada tipo de mensagem existe um tipo de reação. Enquanto que as afirmações feitas pelo emissor não requerem nenhuma mensagem por parte do receptor, as ordens requerem que o mesmo realize o que foi ordenado e as perguntas requerem que o receptor dê uma resposta apropriada.

Para que o emissor obtenha a devida reação por parte do receptor, uma série de fatores extralinguísticos devem ser considerados. Como primeiro elemento podemos citar o nível de cultura. Antes de codificar a mensagem o emissor deve considerar o nível de cultura do receptor, a fim de que sua mensagem seja compreen-

dida e alcance o objetivo por ele pretendido. Outro fator importante é o meio usado pelo emissor para transportar sua mensagem. Caso o emissor esteja utilizando um meio de comunicação de massa, tal como o rádio, a televisão, o jornal, ou outros, o código usado por ele deve diferenciar-se daquele que usaria ao se referir pessoalmente a um receptor ou a um grupo de receptores. Também existem diferenças ao usar um ou outro meio, pois, ao passo que na televisão ele pode utilizar-se de gestos para facilitar a compreensão da mensagem, no rádio ou jornal ele não pode contar com este elemento.

Levando em consideração os pontos discutidos até então, faremos um relato do que deveríamos medir quando testamos a produção oral. Pelas explanações a respeito do que seja a comunicação oral e todos os aspectos que devem ser levados em conta para que a comunicação seja bem sucedida, verificamos que testar a parte oral, assim como a escolha do teste, não são problemas de fáceis soluções.

Tudo que for importante para o ato da fala deve ser examinado quando testamos a habilidade oral, portanto a fluência é o elemento relevante. Dentro da fluência, aparecem os elementos sintáticos, fonológicos e semânticos, revelando a capacidade de comunicar-se do aluno na língua em estudo. Ambos os elementos lingüísticos e extralingüísticos estão envolvidos na

avaliação da comunicação, assim sendo, ambos devem ser transmitidos aos alunos no decorrer da aprendizagem, para que possam utilizá-los corretamente.

2.2. TESTES E SEUS CONCEITOS

Na parte que segue abaixo falaremos sobre testes e conceitos a eles ligados; primeiramente em um plano geral, para depois partirmos para aqueles ligados diretamente à língua. Mais especificamente, àqueles relacionados com a parte oral da língua. Concluiremos a seguinte parte com a descrição do teste N°3: *Instantaneous Evaluation of Speaking Ability*.

2.2.1. Propósito, forma e conveniência de testes

Quando falamos a respeito de testes, algumas perguntas devem ser levadas em consideração. O que é um teste? Quais os diferentes tipos de testes? Para que eles são usados? E quando escolhemos este ou aquele teste outras perguntas devem ser questionadas, assim como: É um teste válido e confiável? Pode ser aplicado diante das circunstâncias sob as quais deve ser aplicado? É muito fácil/longo?

Verificamos que testes são aplicados, no final de um ano, semestre ou até mesmo um mês, por professores que desejam verificar o progresso e o nível de aproveitamento de seus alunos durante aquele período de tempo. Portanto, testes são sempre usados para medir alguma coisa, verificam progresso. No caso do teste oral de língua, ele tenta medir a habilidade oral do aluno para se expressar na língua em questão.

De acordo com MACKEY⁴, testes diferem em três aspectos: propósito, forma e conveniência. Estes três aspectos que serão agora tratados, foram baseados não somente nas idéias de MACKEY, mas também na de outros autores.

a) Propósito

Com relação ao propósito, muitos dos autores consultados mencionaram quatro tipos de propósitos:

- I - testes gerais (*achievement*) - estes tipos de testes são normalmente aplicados no final de cursos e têm como propósito fazer uma avaliação do que o aluno realizou durante o curso, o que ele aprendeu. Os itens deste teste estão baseados em um currículo conhecido.
- II - testes de proficiência - têm como objetivo descobrir o nível de conhecimento que uma pessoa tem em uma determinada língua estrangeira. Testes de pro-

ficiência delimitam o conhecimento para algum propósito e podem variar nas diferentes habilidades.

Testes de proficiência apresentam duas subcategorias de acordo com John CLARK⁵: direto e indireto. Testes diretos tentam refletir a vida real o máximo possível, assim como acontece na entrevista oral em que o examinador e o examinando estão face a face tentando manter um diálogo. Por outro lado, testes indiretos não refletem a vida real. Um exemplo de teste indireto no campo da habilidade oral é a descrição de gravuras.

III - testes de aptidão - medem o grau de facilidade de uma pessoa para aprender certa língua, e não estão baseados no conhecimento da língua.

IV - testes diagnósticos - revelam o que falta ser ensinado ou revisado, com base nos erros cometidos no teste pelos alunos. Após a aplicação de testes diagnósticos uma recuperação pode ser dada.

b) Forma

Como segundo elemento que ocasiona diferença entre testes aparece a forma, na qual diversos aspectos são levados em consideração por MACKEY⁴. Em primeiro lugar devemos verificar qual a habilidade que está sen-

do testada (leitura, escrita, fala, ou compreensão), e como está sendo testada (testes objetivos ou não objetivos). Após esta constatação devemos verificar quais os tipos de perguntas que o teste apresenta, prestando atenção para os seguintes fatos: distribuição, explicações e grau de dificuldade das perguntas. As perguntas podem ser faladas, escritas, através de gravuras, objetos e ações, e ainda individuais ou para o grupo.

c) Conveniência

A terceira diferença entre os testes apresentada por MACKEY é a conveniência. Aqui, teremos que verificar se o teste é adequado para o propósito ao qual será ou está sendo aplicado, como, por exemplo: Testa o que se propõe testar? Qual é o nível de dificuldade? O tempo disponível para responder é suficiente ou demasiado? Pode ser aplicado e corrigido pelo pessoal disponível?

2.2.2. Critérios para comparação com relação à eficiência - confiabilidade, validade e correlação

Existem dois conceitos de fundamental importância para testes. Estes conceitos são a confiabilidade e a validade de um teste.

. Confiabilidade

Dois tipos de confiabilidade são mencionados por Elizabeth Ingram em ALLEN & DAVIES⁶: confiabilidade estável e confiabilidade equivalente.

I - Confiabilidade estável - este tipo de confiabilidade acontece quando obtemos o mesmo resultado, após aplicarmos um certo teste, duas vezes com a mesma amostra. Uma confiabilidade estável perfeita é quase impossível acontecer, pois ela aconteceria quando além de obtermos a mesma ordem entre os indivíduos estes conseguissem a mesma nota. É considerado um bom resultado, quando conseguimos a mesma ordem entre os indivíduos ao aplicarmos o teste pela segunda vez.

II - Confiabilidade equivalente - quando aplicarmos dois testes diferentes, porém que avaliem a mesma habilidade, a um mesmo grupo de pessoas e conseguirmos uma ordem idêntica entre os indivíduos nos dois testes, então este teste apresenta um alto grau de confiabilidade equivalente.

. Validade

No que concerne à validade de testes vemos que ao se propor medir algum elemento, algum aspecto, todo

teste tenta apresentar validade, podendo oferecê-la ou não. Conseqüentemente, um teste oferece validade quando ele realmente mede o que se propõe medir. Tomamos conhecimento de três tipos de validade através E. Ingram em ALLEN & DAVIES⁶ e OLLER⁷, os quais são: validade concorrente, de conteúdo e construtiva.

A validade concorrente sucede quando dois testes que se propõe avaliar a mesma habilidade realmente a avaliam. Obtém-se a validade concorrente através da comparação entre os resultados do teste e uma medida padrão (*criterion measure*) pré-estabelecida. De acordo com E. Ingram: *Criterion measures can be standard examinations, established tests, the ratings obtained from experienced judges on observational samples*⁶ (p. 19). Alguns metodologistas denominam a validade concorrente por validade pragmática.

Validade de conteúdo, como já está explícito pelo nome, diz respeito ao conteúdo. O conteúdo do teste deve conter aquilo que ele se propõe avaliar. Um teste de tradução, por exemplo, deve avaliar somente tradução. Com relação à validade do conteúdo deve-se observar também a validade do currículo, uma vez que testes e currículos mantêm um elo de ligação. Sendo que testes devem basear-se no currículo, este deve também apresentar validade, isto é, deve ensinar o que realmente está descrito nos seus objetivos. Portanto, cur-

rículos e testes estão ambos sujeitos à validade. *The ultimate validity criterion is not a mere requirement for tests alone, but it must be a requirement for the curriculum itself.* OLLER⁸ (p.404).

No que concerne aos currículos devemos observar se eles produzem o que prometem produzir. OLLER⁸ diz que a maioria dos currículos do ensino de língua estrangeira não geram o que prometem, não produzem pessoas que ao final do curso sejam capazes de usar a língua por eles transmitida.

O terceiro tipo conhecido por validade construtiva refere-se ao embasamento teórico dos testes. Todos os testes devem seguir certas exigências teóricas, e quando isto realmente acontece, o teste apresenta validade construtiva. Existem vários conceitos teóricos com respeito a testes, isto é, cada tipo de teste diz pertencer a um certo conceito teórico de acordo com a sua forma, estrutura e conteúdo. Assim que o dizer pertencer dos testes coincide com o realmente pertencer pode-se afirmar que tal teste tem validade construtiva. Por exemplo, quando o teste for classificado de acordo com o propósito como um teste de proficiência ele tem que apresentar os requisitos exigidos em um teste deste tipo. No caso de tal fato realmente acontecer, podemos afirmar que a validade construtiva está presente neste teste, ao passo que quando ele não apre-

sentar os requisitos exigidos pelo teste, a validade construtiva não faz parte deste teste.

2.2.3. Contagem de pontos e aplicação

Faremos uma breve referência à contagem de pontos e à aplicação, por não apresentarem grandes dificuldades. Ao tratar da aplicação e da contagem de pontos, John CLARK⁵ menciona diferentes aspectos.

Aplicação é o processo usado para apresentar o teste ao aluno, podendo ser mecânica ou humana. No caso da aplicação mecânica, utilizam-se aparelhos tais como gravadores, retroprojetores, *slides* e outros.

Contagem de pontos é o processo usado para relacionar as respostas dos alunos com números. Assim como a aplicação, a contagem de pontos pode também ser mecânica ou humana, e pode ainda ser dividida em simultânea ou retardada. É simultânea quando fornece o resultado na hora, e retardada quando é apresentada dias após o teste.

A contagem de pontos retardada e mecânica apresenta algumas vantagens em relação à simultânea e humana. A fita na qual são gravados os testes pode ser ouvida quantas vezes for necessário em caso de dúvidas, e não existe a chance de o examinador favorecer

um ou desfavorecer outro, resultante da simpatia ou antipatia que sente por certos alunos.

2.2.4. Estágios pelos quais passaram os testes relacionados à língua nos últimos anos

- Relação entre testes, objetivos e currículos.

Quando preparamos ou quando escolhemos testes orais de língua, devemos ter sempre em mente dois fatores importantes do ato de comunicação, os quais são: o que dizer, e a pessoa a quem dizer. Portanto, partindo-se dessas duas premissas que são de extremo valor para a comunicação, verificamos que o significado em testes de língua deve realizar um importante papel.

Alan DAVIES, em seu artigo *Language Testing*⁹ menciona o fato de ter havido uma mudança em testes de língua, em especial naqueles relacionados com a parte oral, nos últimos anos. Ele diz que em meados dos anos 60 era dada muita atenção ao som, à parte fonológica; enquanto que por volta dos anos 70 o interesse voltou-se ao significado e comunicação.

De acordo com Spolsky em Alan DAVIES⁹ existiram três estágios no desenvolvimento de testes no decorrer deste século:

1. O teste pré-científico

Neste primeiro estágio os testes eram praticamente de traduções, composições e completar sentenças. No que concerne à validade e confiabilidade de testes, nada se conhece.

2. O teste estrutural-psicométrico

Esta fase caracteriza-se pela objetividade dos testes. O aumento da confiabilidade desse teste deve-se ao fato de ocorrer maior facilidade na parte da avaliação de testes objetivos do que dos subjetivos. Em testes objetivos não existe a possibilidade de o examinador mudar de opinião durante a correção nem de o examinando expressar-se mal. Temos como exemplo desta fase o *TOEFL*.

3. O teste psicossociolingüístico

Apresenta-se como fator principal, nesta fase, o aspecto integrativo, fazendo surgir os testes que procuram avaliar a proficiência das pessoas em sentido global, obtida durante a emissão de uma mensagem.

Nos últimos anos, a preocupação em relação a melhores testes que meçam a habilidade oral da língua tem sido muito grande. Como afirma Robert Lado em SEWARD¹⁰: *testing the ability to speak a foreign language is perhaps the least developed and the least practised*

(skill) in the language testing field (p.76).

Testes de proficiência oral são de grande importância devido à sua íntima relação com a habilidade de fala e devido ao novo enfoque que tem sido dado ao aprendizado da língua, o qual deposita interesse em implantar a habilidade oral nos alunos. Portanto, não há coerência ao submetermos alunos, de cursos onde o novo enfoque esteja sendo aplicado a testes cuja habilidade testada seja outra que não a oral. Concluimos então, que testes deveriam avaliar o que é ensinado durante o curso, isto é, deveriam avaliar os objetivos propostos no início do curso. Podemos afirmar que uma seqüência lógica seria primeiramente verificarmos os objetivos, os quais nos conduzem à realização da aprendizagem, que por sua vez é verificada através dos testes.

Objetivos de um curso pertencem a um currículo. São estabelecidos antes do início do curso, e são os propósitos almejados pelo curso. Eles devem ser claros e razoáveis para as metas almejadas. De acordo com BANATHY & LANGE¹¹ existem objetivos específicos e não-específicos, apresentando-se os não-específicos antes dos específicos. O grau de especificidade dos objetivos vai determinar a especificidade dos testes, isto é, quanto mais específicos forem os objetivos mais específicos deverão ser os testes. E, a partir do prin-

cípio de que os testes são a comprovação de que os objetivos foram alcançados, aqueles deverão ser fiéis aos objetivos.

No prefácio escrito por MACKEY para o livro de JAKOBOVITS¹² ele faz as seguintes asserções: *The language instruction a learner gets(I) is a function of both teaching(T) and method(M), which in turn are functions of curriculum and Language Policy (p.13). The curriculum can determine what, how much and how the learner learns through the methods and materials with which he comes in contact (p.13).* Se o método e ensino estão intimamente ligados ao currículo, testes também estarão ligados ao currículo, uma vez que se relacionam com método e ensino. Já que o currículo determina o que e quanto ensinar, o teste de realização deveria relacionar-se a ele pois mede o que e quanto foi ensinado; o que e quanto o aluno aprendeu.

O currículo pode, por exemplo, delimitar os aspectos ou a habilidade a ser ensinada. Quando isso acontece, os testes deverão ser rigorosamente escolhidos combinando-se com a habilidade que o currículo se propõe enfatizar. No caso do currículo optar pela escrita, os testes deverão avaliar a escrita.

2.2.5. Testes integrativos & testes analíticos

- Participantes & não-participantes

Alan DAVIES nos seus textos *Language Testing*⁹⁻¹³ faz referência a duas divisões de testes: integrativos e analíticos. Os testes ditos integrativos interessam-se pela língua como um todo, isto é, testam a língua como um todo, enquanto que os analíticos testam as estruturas separadamente. Estas estruturas são denominadas de pontos específicos (*discrete points*). Existem, ainda, os testes que são uma combinação dos dois; testam a língua como um todo, apresentando, porém, itens de pontos específicos dentro do todo, dentro do texto.

No entanto, antes de optarmos pelo estilo analítico ou integrativo para a elaboração de um teste, teremos que verificar quais os propósitos e as condições em que o teste será aplicado. Baseando-se nessas condições, mais ênfase será dada ao aspecto analítico ou ao aspecto integrativo. No que concerne a testes orais, levantamos a seguinte pergunta: testes de proficiência oral deveriam conter itens de pontos específicos ou deveriam basear-se em uma avaliação mais integrativa?

Após alguns estudos chegamos à seguinte conclusão: Existem alguns fatores que dão prioridade aos testes integrativos na avaliação da proficiência oral. O

primeiro ponto em favor do uso de testes integrativos é o fato de a língua ser um todo, e não um conjunto de pedaços desligados um do outro. Como segundo argumento aparece o objetivo da língua. Sendo que o objetivo prioritário da língua é a comunicação, é ela que deveria ser testada e não pontos específicos de conhecimento. Existe ainda um outro ponto a ser considerado, o qual diz respeito à generalidade de itens de pontos específicos. Uma pessoa pode conhecer vários itens gramaticais, mas não ser capaz de comunicar-se na língua em questão. Neste sentido testes de pontos específicos podem ser eficazes para outros propósitos que não a competência comunicativa.

Allen DAVIES¹³ cita como exemplos de testes integrativos os seguintes testes: *cloze*, ditado, tradução e entrevista oral. E ainda defende a hipótese de que a confiabilidade e a validade podem aumentar pelo uso de pontos específicos e testes do tipo integrativo, respectivamente. Quanto maior a incidência de itens de pontos específicos maior a confiabilidade, e quanto mais integrativo for o teste maior será a validade.

Para tentarmos chegar a uma conclusão com relação a estes dois aspectos, diremos que testes gramaticais deveriam utilizar itens discretos e testes de comunicação deveriam usar avaliações de aspecto integrativo.

Peter ROBINSON¹⁴ apresenta uma outra divisão para testes da língua oral, a qual ele denomina de teste participante e teste não-participante. No teste participante o examinando constrói suas próprias respostas; ele usa a língua. No teste não-participante ele escolhe suas respostas dentro de uma série de respostas a ele apresentada. Neste último caso ele mostra o conhecimento que tem da língua e não a capacidade que tem de usar a língua. Através da escolha da resposta correta, o examinador não pode garantir que o aluno seria capaz de usar a forma correta se esta não lhe tivesse sido apresentada; ao passo que no teste participante o aluno emite formas corretas e incorretas.

Como a única maneira de sabermos se um indivíduo pode comunicar-se em certa língua é fazê-lo falar, o teste participante parece ser mais adequado para a presente circunstância, que é a habilidade oral em comunicar-se. No entanto, para fazermos um indivíduo falar o teste deve ser motivante e real, em que a motivação está relacionada à realidade. A língua usada no teste deveria ser a da vida real e a escolha de tópicos e perguntas a serem discutidos deve ser de interesse do examinando. Para tanto, é recomendável haver sempre mais do que uma opção. *The role of the oral expression test is to do just that: create a situation in which the student participates freely and at his*

best. Peter ROBINSON¹⁴ (p.266).

2.2.6. Proficiência oral

Como e qual a melhor maneira de testar a proficiência oral é um problema que ainda não obteve uma solução bastante satisfatória. Vemos que o problema tem como raiz a falta de uma definição exata para o que seja a proficiência oral, conduzindo-nos a medir o desconhecido, o que se torna muito difícil, visto que só podemos medir alguma coisa a partir do momento em que sabemos o que seja.

Vários elementos giram em torno de uma definição para a proficiência oral, a qual tem-se tentado definir incluindo os seguintes elementos: velocidade, exatidão gramatical, uso de vocábulos apropriados e exatidão de pronúncia, sem, no entanto, chegar-se a uma decisão. No entanto, não temos quais os requisitos, bem delineados e explícitos que devem ser preenchidos por um indivíduo para que ele seja considerado proficiente.

Outro fator de destaque ao testarmos a proficiência oral é o fato de que raramente um indivíduo será proficiente em todos os setores. De acordo com as suas necessidades, cada indivíduo será mais ou menos

proficiente em diferentes áreas. Estas áreas estão relacionadas com as suas necessidades em comunicar-se; portanto, ao testar a proficiência, seria uma situação ideal se os examinadores tomassem prévio conhecimento das áreas ou setores e níveis sociais em que os examinandos terão que usar a língua.

Concluindo, verificamos que ser proficiente em uma língua é conseguir entender e ser entendido ao comunicar-se, e o nível de proficiência varia de acordo com os interesses e necessidades das pessoas.

2.2.7. Diferentes maneiras de medir a proficiência oral

Ultimamente, tem-se dado muito ênfase aos testes orais, devido ao fato da língua ser primeiramente um elemento de comunicação no qual a parte oral tem prioridade fazendo com que muitos cursos dêem ênfase ao desenvolvimento oral da língua. Essa prioridade cedida à parte oral faz com que os testes sejam também dirigidos a esta área. Por este motivo, e pelo simples fato de o intuito deste trabalho ser a análise de um desses testes orais, faremos nesta parte uma exposição dos diferentes testes existentes para avaliar a parte oral da língua.

I - A entrevista oral

Dentre os vários tipos existentes de testes para medir a proficiência oral a entrevista oral tem sido o mais usado. Embora a entrevista não seja completamente satisfatória e apresente falhas, nenhuma outra maneira mostrou-se superior. Tem-se tentado encontrar uma maneira padrão para testar a produção oral, que desse melhores resultados que a entrevista oral, que fosse mais objetiva, mais rápida, e proporcionasse dados concretos para uma correção menos subjetiva e mais justa.

As limitações atribuídas à entrevista oral seguem abaixo:

1. Subjetividade na avaliação. Os examinadores ao avaliar uma entrevista oral são subjetivos devido ao fato de não haver nenhum critério estabelecido em que se basear. Portanto, podem existir diferenças no resultado de dois examinadores, ou até mesmo entre duas avaliações feitas pelo mesmo examinador com o mesmo aluno em ocasiões diferentes.
2. Consome muito tempo para ser aplicado.
3. Caro. A entrevista oral torna-se cara, pela necessidade de ter examinadores treinados. Em muitos casos esses examinadores são obrigados a ser falantes

nativos da língua em questão, o que além de caro fica mais difícil.

4. Muitas vezes o aluno não tem chance de fazer perguntas como está estabelecido nos padrões de uma entrevista e como ocorre em uma comunicação real. Esta falha ocorre por causa da falta de tempo; existe, muitas vezes, um grande número de alunos para poucos examinadores.
5. Um outro aspecto que poderia ainda ser considerado, como complementar ao item número um, é o humor do examinador. Muitas pessoas deixam-se influenciar por seu humor, deixando a avaliação mais subjetiva ainda.

Apesar de todas as desvantagens da entrevista oral é difícil encontrar uma técnica mais válida para medir a proficiência oral. De acordo com Jones em C. WILDS¹⁵ *the oral interview is a very valid system* (p. 39). JONES¹⁶ ainda afirma que os testes que usam textos, gravuras ou gravações para motivar o aluno não são melhores que a entrevista, uma vez que a resposta do examinando é sempre limitada e não é natural. De acordo com a classificação de testes em diretos e indiretos, a entrevista oral pertence à categoria dos diretos; espelha a vida real.

JONES¹⁶ cita que o Foreign Service Institute (F.

S.I.) e o Central Intelligence Agency (C.I.A.) fazem uso de entrevista oral para medir a proficiência oral dos interessados em fazer carreira nestas entidades. O F.S.I. e a C.I.A. adotaram a entrevista oral por ela ter-se mostrado superior aos outros métodos de avaliação. No caso destas duas entidades, o teste é aplicado em pessoas cujo trabalho depende do saber usar a língua.

Um dos argumentos propostos por aqueles que são contra o uso da entrevista oral é o fato de ela causar ansiedade. No entanto, ao analisarmos mais detalhadamente, vemos que todo e qualquer teste causa ansiedade, especialmente quando notas e salários, como é o caso da C.I.A. e do F.S.I., estão dependendo do resultado do exame.

OLLER², no capítulo onze de seu livro, dá uma explicação detalhada do *Foreign Service Institute Oral Interview*, cujos procedimentos mais importantes estão descritos abaixo.

No teste do F.S.I., dois examinadores testam cada indivíduo por quinze minutos, gravando as entrevistas para futura referência. O treinamento dos examinadores no F.S.I., que é um ponto muito importante, teve uma grande melhora nos últimos anos, pois, existem boas bibliotecas com fitas de testes em todos os níveis, nas línguas mais testadas, para serem usadas pelo novo exa-

minador no período de treinamento. A proficiência máxima da fala, almejada, baseia-se na proficiência que um falante nativo tenha da língua. O exame tem o intuito de ser aplicado, principalmente ao pessoal do governo ligado a assuntos internacionais.

A entrevista determina o nível do indivíduo, com base numa escala de cinco pontos:

- 1) *able to satisfy routine travel needs and minimum courtesy.*
- 2) *able to satisfy routine social demands and limited work requirements.*
- 3) *able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social and professional topics.*
- 4) *able to use the language fluently and accurately on all levels normally pertinent to professional needs.*
- 5) *speaking proficiency equivalent to that of an educated native speaker. (p.320)*

No que se refere aos pontos específicos, o F.S. I. faz uma avaliação separada, ou seja, usa uma outra escala para avaliar os pontos específicos. A escala consiste de: exatidão de pronúncia, gramática, vocabu-

lário, fluência e compreensão. Contudo, o resultado obtido nesta escala não modifica o resultado obtido com a escala anterior; somente adiciona alguns dados. O ponto gramatical recebe o maior peso seguido de vocabulário, compreensão, fluência e intensidade, que recebe o menor peso.

II - Entrevista oral realizada na Universidade de Erlangen

Esforços para melhorar os testes orais de língua foram também realizados na Universidade de Erlangen, em Nürenberg, cujas tentativas foram relatadas por TOWSON¹⁷. Melhoramentos neste setor da faculdade tornaram-se de extrema necessidade, uma vez que esta entidade forma pessoas que seguem a carreira do ensino do inglês. Estes alunos prestam dois exames, sendo que ambos contêm uma parte oral. Pela razão de futuros professores de língua necessitarem ter um ótimo comando da parte oral, muita ênfase é depositada nesta parte.

A forma de teste usada é a entrevista. Uma vez que a intenção é reproduzir a comunicação oral da melhor maneira o aluno é imediatamente deixado à vontade e são evitadas perguntas em que a falta de informação impessa que os alunos respondam adequadamente. Para que não haja nenhuma correlação entre a escrita e a fala,

isto é, para que a compreensão oral não seja influenciada com símbolos visuais, não é apresentado nenhum material escrito junto com a entrevista.

A fim de evitar o problema da avaliação subjetiva, de dois examinadores fornecerem resultados diferentes, estabeleceu-se uma folha de avaliação padrão. Sugeriram-se cinco aspectos a serem julgados:

- a) pronúncia
- b) sintaxe
- c) vocabulário
- d) fluência
- e) compreensão oral

Os alunos são avaliados sob dois aspectos; a habilidade receptiva e a produtiva. Na habilidade receptiva os alunos deveriam entender o inglês como ele é falado, e na habilidade produtiva, dois outros conceitos são aplicados:

- a) compreensibilidade
- b) quase nativo

Para o propósito a que se destinavam os formandos o primeiro foi considerado muito baixo, tornando-se o segundo o ideal a ser desejado.

Problemas a serem solucionados ainda:

- converter os pontos dados para notas;
- valor a ser dado a cada aspecto;
- onde colocar o limite mínimo para ser aprovado.

Como solução para estes e outros problemas, que porventura surgissem, decidiu-se que o examinador marcaria somente os pontos e decidiria a nota final, juntamente com os outros examinadores.

III - Teste oral - Peter ROBINSON

Na opinião de Peter ROBINSON¹⁸ um teste deveria ser dividido em duas partes: uma entrevista que deveria deixar o indivíduo à vontade e seria formada de perguntas pessoais, e uma discussão composta por uma narrativa. A análise, ou seja, a avaliação da produção oral deveria ser padronizada de acordo com os seguintes itens:

- duração do teste;
- número de palavras produzidas pelo indivíduo;
- número de sentenças;
- tipo de sentenças;
- tipo de estrutura gramatical;
- tipo de vocabulário usado;
- pronúncia usada;

- tipo de erro gramatical;
- tipo de erro semântico;
- tipo de erro fonético.

IV - Teste oral - BEARDSMORE & RENKIN

BEARDSMORE & RENKIN¹⁹ sugerem que um teste da parte oral da língua deve fazer uso de diferentes técnicas ao mesmo tempo. O teste oral de inglês proposto por BEARDSMORE & RENKIN pretende fornecer uma apreciação global da qualidade de comunicação. O critério mais importante é a inteligibilidade; é de extrema relevância o indivíduo conseguir fazer-se entender.

O teste oral compõe-se das seguintes técnicas: (a) *strip-cartoon* (técnica que consiste em descrever uma série de gravuras); (b) entrevista; (c) diálogo; e (d) narração sobre um determinado tópico por um minuto.

Usando todas estas variações o indivíduo terá diferentes oportunidades para demonstrar a sua capacidade de comunicar-se na língua pretendida, e o examinador poderá ter mais certeza quanto ao resultado. Terá vários exemplos em que basear a sua avaliação.

V - Bilingual Syntax Measure²

Entre os testes usados para medir bilíngües, o *Bilingual Syntax Measure* (B.S.M.) alcançou uma posição de realce. Tendo sido criado especialmente para crianças ele avalia a habilidade destas em duas ou mais línguas, compondo-se de gravuras coloridas e realmente incentivadoras, a respeito das quais as crianças respondem perguntas sobre o que está acontecendo.

O fato de estarem contando um acontecimento ao examinador, e não somente dando nome aos objetos torna-se motivante para as crianças, e, à medida que as crianças narram um acontecimento, o B.S.M. vai de encontro ao conceito pragmático da língua, no qual como já foi dito a pessoa sente a necessidade de dizer alguma coisa concreta. A série de gravuras apresentadas pelo B.S.M. proporciona um discurso interessante por relatar um acontecimento em série, estando as gravuras entreligadas uma com a outra.

Na contagem de pontos cada tipo de pergunta recebe um determinado valor. Perguntas que podem ser respondidas com o uso de somente um elemento gramatical valem um ponto, as que exigem dois elementos valem dois, e assim por diante. Se tomarmos como exemplo as seguintes respostas *he/didn't want/ to get caught* e *getting/ an orange* verificamos que ao passo que a pri-

meira vale três pontos a segunda vale dois pontos; e no processo de contagem de pontos cada parte da frase correta vale um ponto. São dadas notas para a parte gramatical separada da parte semântica, sendo que a colocação exata da parte semântica é mais importante do que a sintaxe. Sentenças cuja parte gramatical está completamente correta, mas cujo significado não se relaciona com o contexto não valem nenhum ponto. Portanto, testes, de acordo com o B.S.M., devem dar mais importância ao significado do que à sintaxe para serem mais válidos e confiáveis.

VI - Ilyin Oral Interview²

Existem outros testes em uso, que também utilizam gravuras, porém conduzindo para uma entrevista mais dirigida. Um destes testes é o *Ilyin Oral Interview*, criado em 1976 por Donna Ilyin em São Francisco.

As gravuras apresentadas neste teste estão em uma ordem cronológica e são representações das atividades de uma pessoa durante o dia. Donna Ilyin elaborou duas séries de gravuras, para que possam ser aplicadas em qualquer dia da semana, de acordo com o dia do teste. Existem também duas versões, sendo uma de 50 itens para os alunos em nível elementar, e outra composta de 30 itens para os níveis mais altos. Uma gran-

de variedade de abordagens pode ser usada com esta técnica. O examinador tanto pode fazer perguntas relacionando-se somente à ordem cronológica apresentada, como pode fazer perguntas a respeito de uma só gravura e estender-se para outras perguntas que não estejam relacionadas com as gravuras, mas tratem do mesmo assunto. Outra alternativa seria utilizar a sequência de gravuras para fazer com que o examinando conte uma história sobre o que ele está vendo.

VII - Oral Communication - John UPSHUR²

O outro teste que é tido como altamente estruturado é o teste *Oral Communication* desenvolvido por John UPSHUR na Universidade de Michigan. Da mesma forma que os anteriores, este também adota gravuras, porém as gravuras não representam uma sequência significativa, e, por não estarem em uma sequência significativa, fazem com que o teste não obedeça ao critério pragmático do ato de fala. Este problema seria solucionado com a organização de uma sequência entre os itens do teste.

O teste constitui-se de conjuntos da mesma figura básica, apresentando diferenças em alguns pontos. Como exemplo, podemos citar duas gravuras com a mesma menina. Na primeira ela tem cabelos longos e está en-

trando em uma sala, e na segunda ela tem cabelos curtos e está saindo de uma sala.

Vários conjuntos de quatro gravuras são dados para ambos, examinador e examinando. Uma terceira pessoa aponta ao examinando uma gravura em cada conjunto de quatro, a fim de que seja descrita ao examinador. As quatro figuras de cada conjunto são colocadas em ordem diferente em frente do examinador e examinando; ocorrendo que, para cada gravura que o examinador descobrir qual seja através da descrição do examinando, este recebe um ponto.

Com relação a este teste, o fator tempo fez uma importante revelação. Quando o tempo não era limitado não havia diferença entre os alunos bem proficientes e os mais ou menos proficientes, ao passo que, quando foi imposto um tempo limite, houve discriminação entre os dois níveis de alunos.

VIII - Narrativa²

Uma variação usada com crianças que pode facilmente ser adaptada para adultos é a narrativa. Um dos modelos de narrativa é recontar uma estória. Inicialmente, uma pessoa faz algumas perguntas à criança para deixá-la à vontade, e então conta-lhe uma estória que ela terá que recontar à outra pessoa que não ouviu a

estória.

Aqui, existe o problema de como avaliar correta e justamente conteúdos diferentes, como dar notas a diferentes conteúdos e no que basear-se. OLLER apresenta, como solução, a fixação de uma escala predeterminada.

O outro modelo consiste em dar à criança um conjunto de vários e diferentes materiais e pedir-lhe que forme alguma coisa; depois de algum tempo um examinador pede que a criança lhe diga o que fez e como fez. O tempo estabelecido para a realização do objeto e para a narração pode variar.

IX - Integração entre testes e habilidades

Os testes relatados até o presente estágio foram todos criados com uma intenção específica, quer dizer, foram criados para testar a comunicação oral. Existem testes que não foram criados para avaliar especificamente uma determinada habilidade, podendo por este motivo ser usados para diferentes propósitos, para avaliar diferentes habilidades, isto é, testes usados para medir outras habilidades que podem ser usados na avaliação da comunicação.

Uma teoria defendida por vários metodologistas é que para avaliar a proficiência oral em uma determi-

nada língua, pode-se utilizar meios indiretos.

a) Compreensão oral

Pode-se dizer qual é o nível de comunicação de um indivíduo em uma determinada língua, através de testes sobre a compreensão oral. *An accurate indication of the student's speaking ability in a foreign language can be derived from his auditory comprehension ability in foreign language* B. H. SEWARD¹⁰ (p.77). No entanto, podemos dizer que a asserção acima deixa dúvidas, à medida em que nem sempre o reconhecimento correto implica a correta produção.

Testes de compreensão oral compreendem normalmente a leitura de um texto a respeito do qual são feitas perguntas do tipo de múltipla escolha. Durante a leitura do texto os alunos não devem tomar nota, e a duração da leitura não deve exceder 45 segundos a fim de evitar problemas de dificuldade de concentração.

GROOT²⁰, no seu texto *Testing Communicative Competence in Listening Comprehension*, apresenta algumas variáveis que podem ser usadas em um teste de compreensão oral para ajudar o aluno. Uma das variáveis é fazer com que os alunos escutem as perguntas antes da leitura do texto, o que fará com que os alunos escutem seletivamente o texto. Uma segunda possibilidade é fazer as perguntas na língua materna.

Uma boa técnica para separar os maus ouvintes dos bons é o uso de elementos *echoic*, os quais são palavras tiradas das passagens e usadas nas alternativas das perguntas de múltipla escolha. Os maus ouvintes obviamente assinalaram as alternativas que contiverem estes elementos.

b) Testes objetivos

Testes objetivos para avaliar a proficiência de uma língua em geral podem também ser usados para medir exclusivamente a parte oral. A conclusão alcançada com o resultado destes testes objetivos foi que eles fornecem um resultado mais seguro da parte oral do indivíduo do que as avaliações dos professores. Para testar a comunicação oral, como para qualquer outro aspecto não se deve confiar somente em um resultado. Por este motivo mais do que um teste objetivo é usado. Verificado isto, dizemos que as avaliações feitas pelos professores devem ser usadas juntamente com outras avaliações.

c) Leitura

A leitura em voz alta também pode ser aplicada para avaliar a habilidade oral de algumas pessoas. De acordo com OLLER², ela é útil para aqueles que lêem bem em alguma outra língua e que vão provavelmente se tor-

nar literatos nessa língua. Uma das maneiras de usar a leitura oral para avaliar a fala seria cronometrar o tempo enquanto a pessoa lê.

A leitura em voz alta como uma maneira de avaliar a fala é objetada por aqueles que dizem ser ela muito mais fácil do que falar. Muitas vezes um aluno seria capaz de ler um texto mas não seria capaz de usar as estruturas e palavras contidas neste texto enquanto falasse. Por outro lado, aqueles que são a favor se contrapõem a essa objeção dizendo que *a foreign language learner, for example, who cannot carry on a simple conversation fluently will probably not be able to read a somewhat more complex text with fluency either* OLLER² (p.328). Uma outra dúvida a ser observada é a de que se um aluno estrangeiro é capaz de ler um texto fluentemente sem entender o significado. Se ele não é capaz, quer dizer, se ele precisa entender o que está lendo para fazê-lo fluentemente, então este torna-se um outro ponto a favor da leitura oral ser usada para medir a proficiência da fala.

d) *Oral cloze*

Um dos procedimentos menos usados para avaliar a habilidade oral é o *oral cloze*, que, no entanto, tem sido altamente usado para medir outras habilidades. Embora os testes orais do tipo *cloze* sejam baseados nos

escritos, eles apresentam algumas diferenças. Enquanto que nos testes escritos o aluno recebe o texto de ambos os lados da lacuna, nos textos orais ele recebe somente a parte que precede o que ele deve completar. Para que não se tornem muito difíceis oferece-se como alternativa a leitura do texto completo uma ou duas vezes.

Testes orais do tipo *cloze* podem ser aplicados de diversas maneiras. A primeira possibilidade é fazer intervalos iguais de supressão, pôde-se suprimir, por exemplo, cada quinta palavra de um texto. Outra possibilidade é darmos ao aluno somente as oito ou dez primeiras palavras da frase e ele tem que completar com a palavra ou as palavras seguintes. Após ele completar o examinador lê o texto seguinte repetindo o anterior. Em vez de completar somente com palavras, o examinador pode pedir para completar com uma sentença. O teste pode ser feito separando sentença por sentença, ou com o texto todo, o que se tornará cansativo devido ao número de vezes que o material precedente terá que ser relido. Para que o teste possa ser aplicado para um grupo de alunos ao mesmo tempo, as respostas devem ser gravadas ou escritas. No caso de se escolher a segunda alternativa fica eliminada a repetição do material correto.

Após termos examinado e estudado várias manei-

ras de testar a habilidade oral concluímos que, apesar de existir um número razoável de testes, nenhum dos métodos até então existentes é satisfatório para ser usado em qualquer situação. O método a ser usado varia de acordo com a situação, isto é, situações diferentes exigem tipos de testes diferentes; e o tipo de teste por sua vez está relacionado com a maneira de ser transmitida a matéria, com o tipo de aula que é ministrado. Para uma turma que tem treinado a parte oral através de descrições de figuras e objetos, nada mais justo do que usarmos gravuras para testá-la.

Como já foi mencionado acima que não existe um método exemplar, resolvemos escolher a entrevista oral realizada por um falante nativo para fazermos a comparação com o novo método N73. Portanto, até que se prove o contrário, a entrevista oral foi determinada por nós como o método mais confiável dentre os existentes. Usamos como critério nesta escolha os fatos de ser a entrevista oral a técnica mais fiel ao diálogo humano e de estar sendo sempre treinada em sala de aula. O professor pergunta e o aluno responde, e o aluno pergunta a aluno.

Uma vez que o teste com o qual vamos compará-la avalia a habilidade oral através de um discurso oral, não seria razoável usarmos, por exemplo, uma das técnicas indiretas tão comentadas por OLLER. Além disto,

achamos que, no que concerne a testes, a integração de habilidades não fornece bons e fiéis resultados. Visto que um grande número de alunos com que tratamos apresenta maior facilidade em escrever, ler e até mesmo entender do que falar durante o período de aula, não obteríamos um resultado fiel da sua habilidade da fala caso usássemos uma técnica indireta; como, por exemplo, um teste de compreensão oral.

2.3. TEST N73: INSTANTANEOUS EVALUATION OF SPEAKING ABILITY

Para finalizarmos esta parte relativa a testes será feito um relato de como funciona o teste N73: *Instantaneous Evaluation of Speaking Ability*, proposto por Nicolas FERGUSON. Como uma introdução para o seu teste, FERGUSON menciona o fato de que a maioria dos alunos não vê nenhum motivo para estudar o que não é testado, portanto os cursos que tiverem como objetivo desenvolver a habilidade oral precisam de uma avaliação também oral.

Nicolas FERGUSON diz que o primeiro passo para a construção de N73 foi definir uma maneira padronizada de medir a expressão oral. A construção do teste foi

baseada em uma teoria matemática da comunicação, conhecida como *stochastic process*. É um sistema que produz uma mensagem, escolhendo símbolo por símbolo, os quais são escolhidos de acordo com certas probabilidades. *Stochastic processes can be defined which produce a text consisting of a sequence of words* (N. FERGUSON²¹ p.341). Estas seqüências de palavras tentam assemelhar-se com o inglês, sendo que esta semelhança aumenta à medida que as estruturas mais avançadas aparecem. É proposto como exemplo da construção das estruturas o seguinte:

Ele afirma que se obtém uma aproximação de primeira ordem do inglês escolhendo-se palavras sucessivas independentemente. Pode-se chegar à primeira ordem ao abrir-se qualquer livro em qualquer página e pegar a primeira palavra em cada linha sucessivamente. FERGUSON cita como exemplo o seguinte: *It lack to quick through excluded called showed who for back main brick ...* (p.341).

Uma aproximação da segunda ordem, forma o que FERGUSON chama de *di-gram structure*. Para obter-se uma aproximação de segunda ordem escolhe-se uma palavra em qualquer página de um livro; a seguir, pega-se um segundo livro e procura-se um exemplo da mesma palavra escrevendo a palavra que a segue. Volta-se para o primeiro livro, procura-se um exemplo da segunda palavra

e escreve-se a palavra que a segue. Prosseguindo desta maneira, forma-se um exemplo como o apresentado por FERGUSON: *You know a time before he did the phone on to blame the lift...* (p.341).

Após *di-gram structure* pode-se obter *tri-gram structure* através de uma aproximação de terceira ordem. Para a formação de *tri-gram structure* é necessário escolher duas palavras consecutivas. A seguir encontra-se outro exemplo das duas palavras e copia-se a palavra que as segue. Pega-se então o último par de palavras e volta-se para o primeiro livro, para continuar o processo.

Ex.: *The first boy that came when I asked him to take a photograph of a person who knows whether he can do the tomatoes...* (p.341)

Este processo de formação de *N-gram structures* continua através de *tetra-gram*, *penta-gram structure* e assim por diante.

O exemplo citado acima de uma estrutura de primeira aproximação resulta em uma *mono-gram structure*, a qual por sua vez produz nada mais do que simples palavras. A estrutura *di-gram* produz grupos de intensidade, e a estrutura *tri-gram* produz grupos tonais. Nicolas FERGUSON afirma que um grupo de intensidade na parte oral é composto por uma sílaba acentuada rodeada por

uma ou várias sílabas não acentuadas (p.342). Grupos de intensidade podem ser simples palavras tais como:

sometimes, exceptionally.

Ou podem ser constituídos por mais de uma palavra:

He's a doctor.

I can do it.

Com relação a grupo tonal, FERGUSON usa o conceito elaborado por Halliday em 1967: *a tone group is composed of one stress group with primary stress together with other less prominent stress groups round it. It gives a basic unit of information* (p.342). O seguinte texto está dividido em grupos tonais.

The children get up at 8 o'clock/and their mother prepares breakfast for them/their father goes to work after breakfast/and they go to school/their mother goes shopping/and then makes lunch at twelve. (p.342)

No exemplo acima, cada sentença corresponde a um grupo tonal. Porém, existem certos casos, como as quatro situações relatadas abaixo, que apresentam casos especiais com relação ao grupo tonal:

1. Complementos que ocorram antes do sujeito sempre formam um grupo tonal separado.

After breakfast/he went to the office.

At 8 o'clock/the children get up.

2. Uma sentença relatora seguida por uma sentença relatada formam um único grupo tonal.

He thinks that you ought to take it.

3. Uma sentença condicional seguida por uma sentença condicionada formam um único grupo tonal.

I'll pay if you come to the cinema.

4. Sentenças relativas formam um só grupo tonal.

Here's the one I wanted to buy.

Os grupos de intensidade e grupos tonais, como foi demonstrado acima, não são determinados pela entonação do falante, mas estão, sim, sendo determinados com bases nas definições de unidades de língua. Estes esclarecimentos se fazem necessários, devido ao fato de eles servirem de base para a avaliação da fluência e exatidão, as quais são os dois elementos medidos com o teste N73.

FLUÊNCIA

A fluência é obtida através do número de grupos tonais por minuto.

$$\frac{T}{t} = \frac{\text{total de grupos tonais}}{\text{tempo total em minutos}}$$

O examinador tem que simplesmente contar os grupos tonais em um determinado tempo.

EXATIDÃO

A exatidão é medida de acordo com o tipo de erro cometido em cada grupo tonal. Diferentes tipos de grupos tonais terão diferentes pesos. Para exemplificar, FERGUSON apresenta o seguinte:

//He goes/to the cinema/on Thursday.//

Este é um grupo tonal composto por três grupos de intensidade. Se um aluno falar:

//He go/ to the cinema/on Thursday.//

ele terá feito um erro de formação no nível mais baixo. O *di-gram* *he go* está incorreto, estando somente correto em um nível *mono-gram*, ou seja, da palavra. Quando isto acontece, este grupo tonal vai ser classificado como aproximação sintática de primeira ordem ou S1.

Se um aluno falar:

//He goes/in the cinema/on Thursday.//

ele formou estruturas *di-gram* corretas, uma vez que todos os grupos de intensidade estão corretos individualmente. Contudo, o grupo tonal *tri-gram* está incorreto, pois, deveria ser: *He goes to the cinema on*

Thursday. O presente caso é classificado como uma aproximação sintática de segunda ordem, S2.

Uma aproximação de terceira ordem, S3, seria a seguinte:

//Last week/he goes to the cinema on Thursday.//

O segundo grupo tonal - *he goes to the cinema on Thursday* - está correto num nível *tri-gram*, mas está fora de contexto.

Um grupo tonal que esteja completamente correto é Sc.

Para calcular o grau de exatidão, FERGUSON pesa cada grupo tonal como segue:

Sc 1

S3 1/2

S2 1/4

S1 0

A exatidão é fornecida por:

$$\frac{EI}{ET} = \frac{\text{total weighting (peso total)}}{\text{total tone groups (total de grupos tonais)}}$$

A maneira para avaliar a exatidão é a mesma tanto para a parte sintática como para a parte semântica. Um grupo tonal que apresente neologismo vai ter o peso

correspondente a S1, um grupo tonal com uma palavra fora de contexto valerá S2, e quando uma palavra ocasiona que um grupo tonal se torne irrelevante em relação a outro terá o peso de S3.

FERGUSON declarou que, ao começar a elaboração de N73, ele havia pensado em incluir mais seis pontos para avaliação além de fluência e exatidão. Estes outros pontos eram os seguintes: hesitação, complexidade sintática, entonação e ritmo, critério fonêmico, critério fonético e estilo. No entanto, estes pontos foram eliminados devido à dificuldade de haver um acordo com relação às opiniões dos examinadores. As conclusões obtidas por eles tinham um alto grau de variabilidade.

Existem dois métodos para se aplicar o teste N73:

- caneta e papel
- através de computador
- Caneta e papel

É necessário que o examinador tenha uma folha de papel na qual estejam escritos os símbolos S1, S2, S3 e Sc. À medida que cada grupo tonal é dito pela pessoa que está sendo examinada, o examinador coloca um traço ao lado da estrutura correspondente. No caso de ele ter emitido um grupo tonal correto, um traço seria

colocado ao lado de Sc. Após o término, o examinador aplica, então, as fórmulas:

$$\text{Fluência} = \frac{\text{total de grupos tonais}}{\text{total de tempo em minutos}}$$

$$\text{Exatidão} = \frac{\text{peso total}}{\text{total de grupos tonais}}$$

- Método através do Computador

Ao mesmo tempo em que este sistema é mais rápido, ele é mais caro. Para aplicação deste sistema é necessário um pequeno computador digital, sendo que o examinador terá em suas mãos um controle remoto com seis chaves: Start, Pause, Sc, S3, S2, S1. Ao começar ele apertará a chave *start*. Então, para cada grupo tonal ele apertará a chave S correspondente; e apertará a *pause* quando for necessário interromper. O resultado de cada indivíduo para fluência e exatidão é obtido simultânea e digitalmente, a cada segundo.

2.4. APLICAÇÕES DOS CONCEITOS AO TESTE N73

Resta-nos ainda por ser feita uma comparação en-

tre a teoria descrita no início do presente capítulo e o teste N73. Tentaremos demonstrar quais os conceitos que são válidos para o teste, isto é, quais das teorias apresentadas se adaptam ao teste.

Concluimos que o teste N73 é uma modificação da entrevista oral, em que somente a produção é testada. Portanto, está lhe faltando um dos pontos muito importantes para a comunicação que é o entendimento, a compreensão dos enunciados. Por outro lado, este teste não apresenta algumas das desvantagens da entrevista oral. Enquanto que a entrevista oral não tem no que basear a sua avaliação, o teste N73 apresenta certas regras para avaliar a fluência e exatidão, sendo possível alcançar resultados exatos. A outra vantagem é o fato de ele não ser caro, no caso da versão *caneta e papel*, e não ser difícil treinar examinadores.

Com relação aos diferentes tipos de testes existentes, N73 pertence à categoria de proficiência. Tem como propósito determinar o nível de um indivíduo na habilidade da fala, e, assim como todos os testes de proficiência, não está baseado em nenhum *curriculum*. Dentro das duas subcategorias de testes de proficiência, direto e indireto, o presente teste pode ser considerado um teste indireto, isto porque ele não é um reflexo verdadeiro do diálogo na vida real. O examinando não faz e não responde perguntas, somente fala a

respeito de um determinado tópic. Esta última afirmação leva-nos a outra conclusão: N73 é um teste participante; o indivíduo mostra a sua capacidade de usar a língua, usando formas corretas e incorretas.

Tentando comparar o que foi dito a respeito de testes analíticos e integrativos, N73 é uma combinação dos dois. Testa a estrutura da língua através de um texto emitido pelo aluno e testa a proficiência levando em consideração os diferentes tipos de erros com relação à estrutura.

Ao compararmos o que LEESON afirma a respeito de fluência e velocidade, e a maneira como FERGUSON mede fluência, verificaremos que eles mantêm idéias completamente opostas. LEESON afirma que fluência não é uma questão de velocidade, e FERGUSON obtém a fluência contando o número de grupos tonais em um determinado espaço de tempo. No entanto, se considerarmos o fato de que N73 está também exigindo dos examinandos a habilidade de manipular as estruturas corretamente, a relação fluência versus velocidade não se torna tão relevante neste caso, não se relaciona com a velocidade em responder. Um outro fato, que deve ainda ser levado em conta é que falantes nativos, quando estão emitindo discursos, necessitam fazer uso de um número mínimo de sentenças, portanto em dois minutos um número X deve ser exigido.

A administração e a avaliação de N73 vão depender do método escolhido. No caso de aplicarmos o método *caneta e papel*, tanto a administração como a avaliação serão humanos, sendo que a avaliação pode ser tanto simultânea como retardada. Mas por outro lado, se usarmos um computador, a administração será humana e a avaliação será mecânica e simultânea.

NOTAS

¹LEESON, Richard. Fluency and Language Teaching. London, Longman, 1975. 202 p.

²OLLER Jr., John W. Tests of Productive Oral Communication. In: _____. Language Tests at School. London, Longman, 1979. p.303-55.

³HERRIOT, Peter. An Introduction to the Psychology of Language. 2.ed. London, Methuen & Co. Ltd, 1971. 189 p.

⁴MACKEY, W.F. The Measurement of Language Learning. In: _____. Language Teaching Analysis. London, Longman, 1966. p.403-416.

⁵CLARK, John L.D. Theoretical and Technical Considerations in Oral Proficiency Testing. In: JONES, R.L. & SPOLSKY, B. Testing Language Proficiency. USA, Center for Applied Linguistics, 1975. p.10-28.

⁶ALLEN, J.P.B. & DAVIES, A. Testing and Experimental Methods. London, Oxford University Press, 1977. 4 v.

⁷OLLER Jr., John W. Discrete Point, Integrative, or Pragmatic Tests. In: _____. Language Tests at School. London, Longman, 1979. p.36-69.

⁸_____. The ultimate problem of test validity. In: _____. Language Tests at School. London, Longman, 1979. p. 403-408.

⁹DAVIES, Alan. Language Testing. Language Teaching & Linguistics Abstracts. London, 11(3):145-59, jul.1978.

¹⁰SEWARD, B.H. Measuring Oral Production in E.F.L. English Language Teaching Journal. London, 28(1):76-80, nov.1973.

¹¹BANARHY, B. & LANGE, D.L. A Design for Foreign Language Curriculum. Lexington, D.C.Heath, 1972. 128 p.

¹²MACKEY, W.F. Foreword. In: JAKOBOVITS, L.A. A Psycholinguistic Analysis of the Issues. Mass, Newbury House, 1971. p.7-13.

¹³DAVIES, A. Language Testing - Part II. Language Teaching & Linguistics Abstracts. London, 11(4):215-26, oct.1978.

¹⁴ROBINSON, Peter. Oral Expression Tests: 2. English Language Teaching. London, 25(3):260-66, jun.1971.

¹⁵WILDS, C.P. The Oral Interview Test. In: JONES, R. L. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.29-44.

¹⁶JONES, R.L. Testing Language Proficiency in the United States Government. In: _____. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.1-9.

¹⁷TOWNSON, M. Testing Oral Skills at University Level. English Language Teaching. London, 27(2):199-205, feb.1973.

¹⁸ROBINSON, P. Oral Expression Tests: 2. English Language Teaching. London, 25(3):260-66, jun.1971.

¹⁹BEARDSMORE, H.B. & RENKIN, A. A Test of Spoken English. IRAL, Alemanha, 19(1):1-11, fev.1971.

²⁰GROOT, P.J.M. Testing Communicative in Listening Comprehension. In: JONES, R.L. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.45-58.

²¹FERGUSON, N. Test N73. Instantaneous Evaluation of Speaking Ability. IRAL, Alemanha, 26(4):340-49, nov. 1978.

3. O EXPERIMENTO

No presente capítulo, será enfim relatado quais foram os passos tomados para a realização do experimento.

O intento do presente trabalho foi medir a confiabilidade do teste N73, o qual teve um tempo de dois minutos, em relação a uma avaliação mais extensa feita por um falante nativo da língua inglesa. O falante nativo que cooperou para a realização do presente trabalho está devidamente qualificado para fazer tais avaliações uma vez que ele tem feito parte da banca examinadora de testes orais no exame de Cambridge por 13 anos. Tendo ele recebido as fitas gravadas por mim, não houve nenhuma relação entre os alunos e o falante nativo, ocasionando uma avaliação pura e simplesmente da amostra contida na fita. As fitas mais extensas avaliadas pelo falante nativo englobaram um tempo de cinco minutos. Ao relacionarmos 2 minutos com 5 minutos em porcentagem, verificamos que 5 minutos são 150% a mais do que 2 minutos; sendo que, de acordo com a opi-

nião do falante nativo e os propósitos do trabalho, cinco minutos seriam o suficiente.

. Material

Foram coletadas amostras de 40 alunos matriculados nos 5ºs. anos elementares e adiantados da Sociedade Brasileira de Cultura Inglesa para a comparação dos dois testes; para melhor colocação do que compreende o 5º ano em tal estabelecimento esclarecemos que transformando em horas/aula, o 5º elementar engloba um total de 432 horas/aula e o 5º adiantado 480 horas/aula. O grupo compunha-se de homens e mulheres de diferentes idades, sendo que todos eles ofereceram-se voluntariamente para serem testados. As amostras foram coletadas no começo do mês de novembro de 1979, e as amostras que foram avaliadas pelo falante nativo foram gravadas em fitas diferentes das do teste N73.

. Condições

As amostras foram gravadas em uma sala da Sociedade Brasileira de Cultura Inglesa; portanto as condições do lugar para obter uma boa gravação não eram bem as desejadas, por haver interferência das outras salas de aula e interferência de fora. Utilizamos para as gravações um gravador minicassete. Foram gravadas

ambas as amostras, ou seja, as amostras que foram analisadas através do teste N73 e as que foram analisadas pelo falante nativo. Uma vez que seria a primeira vez que aplicaríamos o teste, decidimos gravar as amostras para uma necessária revisão e por ser a primeira experiência desse tipo.

. Procedimento

As gravações para o teste N73 e para o falante nativo dos 40 alunos foram realizadas em três ocasiões. Em cada uma das ocasiões, a administração do teste realizou-se da seguinte maneira:

Primeiramente, conversávamos com cada professora e explicávamos o propósito das gravações. Então, ela apresentava-nos aos alunos e deixava-nos conversar com eles, e os que desejassem colaborar eram enviados para a sala mencionada acima. Para que o desenvolvimento da aula não fosse interrompido, dois alunos eram enviados de cada vez, e quando o primeiro voltasse um terceiro aluno viria colaborar.

Para as gravações do teste N73, foram elaborados dois grupos de perguntas, contendo 4 tópicos cada um. Os alunos escolhiam um deles e falavam por dois minutos. Os dois grupos eram compostos da seguinte maneira:

- Grupo A.

1. - What do you do at the weekends?
2. - What do you do every morning?
3. - What is or what are your favorite seasons? Why?
4. - What's your opinion about television? Advantages and disadvantages.

- Grupo B.

1. - What do you do in your leisure time?
2. - What do you do in the afternoon?
3. - Where would you like to spend your next holiday? Why?
4. - What's your opinion about life in the country contrasted with that in the city?

Alguns dos alunos mudaram e incluíram elementos extras em certos tópicos.

Para as entrevistas que foram avaliadas pelo falante nativo, o aluno respondeu perguntas sobre gravuras. Havia quatro gravuras diferentes, sendo que somente uma era usada para cada aluno. As gravuras utilizadas foram retiradas do livro *English in Mind*¹, e descreviam as seguintes cenas:

1. Mostra a fotografia de uma alfândega em um aeroporto, onde aparece um casal que é interrogado pelo oficial.
2. É uma gravura de uma sala de aula, na qual os alunos estão dispostos em círculo. Eles estão provavelmente tendo uma aula de conversação.
3. Uma feira-livre.
4. Fotografia em dois planos. No primeiro plano aparece um parque com algumas pessoas, e no segundo plano há altos edifícios.

A maneira através da qual se procederam as gravações foi a descrita abaixo. Entregávamos a cada um dos dois primeiros alunos um conjunto com quatro perguntas e lhes dávamos algum tempo para organizarem suas idéias a respeito do que iriam falar. Os outros alunos recebiam os tópicos logo que chegavam e tinham tempo para preparar os seus dois minutos de fala enquanto o aluno anterior estava sendo entrevistado. As medidas tomadas para a coleta do material foram as seguintes: logo que o aluno entrava na sala ele dizia o tópico que havia escolhido e começava a falar durante os dois minutos. Durante este tempo, ele não era interrompido e nos avisávamos quando os dois minutos haviam terminado. Terminada esta parte, nos mostrávamos uma das

gravuras e fazíamos perguntas a respeito dela por cinco minutos. Em alguns casos, foram feitas perguntas fora do assunto da gravura além das perguntas relativas às gravuras. Estas fitas foram então entregues ao falante nativo, o qual baseou a sua avaliação na capacidade demonstrada pelo indivíduo de comunicar-se.

Para avaliar as falas de 2 minutos, foi usado o método N73 de *caneta e papel*.

N O T A

¹TTOFI, C. & CREED, T.S. English in Mind. London,
Macmillan, 1974. 194 p.

4. ANÁLISE ESTATÍSTICA DOS ESCORES

- COMPARAÇÃO DOS TESTES

Nesta parte, nós apresentaremos os resultados obtidos pelos alunos nas entrevistas avaliadas pelo falante nativo e a sua comparação com os resultados obtidos com o teste N73. Uma vez que com a aplicação deste teste obtivemos duas variáveis em escalas diferentes da usada pelo falante nativo, e também diferentes entre si, foi feita a transformação de escalas e a média das duas variáveis, preparando-a para a correlação entre as notas obtidas pelos dois processos. Para tal correlação usaremos o coeficiente de correlação por postos de Kendall T (TAU). Após verificada a existência ou não da correlação entre as duas notas, aplicar-se-á o teste de significância o qual conduzirá a uma conclusão.

4.1. RESULTADO DO FALANTE NATIVO

Embora a escala adotada nas entrevistas avalia-

das pelo falante nativo (FN) tenha sido de 0 a 10, as suas notas variaram entre 4 e 7, sendo que somente os alunos que obtiveram entre 5 e 7 seriam aprovados. As notas com seus relativos conceitos são as seguintes: 4 - fraco; 5 - regular; 6 - razoavelmente bom; 7 - bom. Os resultados alcançados por cada aluno nas entrevistas avaliadas pelo falante nativo estão demonstrados no Quadro abaixo:

Aluno	Nota
1	6
2	5
3	4
4	5
5	6
6	5
7	6
8	5
9	6
10	5
11	4
12	7
13	6
14	6
15	5
16	5
17	4

(continua)

Aluno	Nota
18	5
19	4
20	4
21	4
22	5
23	5
24	4
25	5
26	4
27	5
28	4
29	5
30	4
31	4
32	4
33	5
34	5
35	6
36	4
37	5
38	5
39	5
40	6

4.2. RESULTADO DO TESTE N73

Com a aplicação do teste N73 cada aluno recebeu duas notas, uma para fluência e outra para exatidão,

sendo que as notas de fluência variam de 0-20 e as notas para exatidão variam de 0 a 1. Uma vez que o teste N73 fornece duas notas, foi necessário fazer uma média para haver possibilidade de uma comparação com as notas obtidas pelo falante nativo. Isto só foi possível após uma transformação das notas, ou seja, fluência e exatidão, para a mesma escala usada pelo falante nativo.

Antes da apresentação do Quadro seguinte, o qual contém as notas obtidas no teste N73, as mesmas notas transformadas e finalmente a média obtida, se fez necessária uma explicação de como foi feita a transformação das escalas, e no que nos baseamos para afirmar que a fluência varia de 0 a 20, e a exatidão de 0-1.

4.2.1. Variações de Fluência e Exatidão - Transformação de Escalas

Para fazermos uma escala de medidas para fluência, tomou-se como limite, para o valor máximo, um indivíduo que fale um total de 20 sentenças por minuto. Como o teste foi efetuado com um tempo de fala de 2 minutos, obtivemos para um escore máximo de fluência o seguinte:

$$\text{Fluência} = \frac{T}{t} = \frac{40}{2} = 20,$$

ou seja, o campo de variação da fluência é de 20 (máximo) até 0 (zero).

$$F = (0 \leftrightarrow 20)$$

Para se obter a exatidão, como já foi mencionado, leva-se em consideração o peso de cada tipo de sentença, ou seja, $S1 = 0$, $S2 = 0,25$, $S3 = 0,50$ e $Sc = 1$. Desta maneira, podemos observar que para um aluno que não tenha tido nenhum erro de exatidão, a sua nota será 1:

$$E = \frac{\text{Total peso}}{\text{Total de sentença}} = \frac{S1 + S2 + S3 + Sc}{20} = \frac{20}{20} = 1$$

Obs.: Neste caso $S1 = 0$, $S2 = 0$, $S3 = 0$, $Sc = 1$ para todas as vinte sentenças.

$$E = (0 \leftrightarrow 1)$$

Tendo-se o campo de variação das duas notas, podemos levá-las para a mesma escala, e após isto obtermos a média aritmética entre ambas e assim compará-las finalmente com as notas do N73:

$$\begin{array}{ccc}
 F = (0 \leftrightarrow 20) & & \\
 \left. \begin{array}{l} \\ \\ E = (0 \leftrightarrow 1) \end{array} \right\} M = N73 (x) & & NFN (y) = (0 \leftrightarrow 10) \\
 & \uparrow \quad \text{correlação por Kendall} \quad \uparrow &
 \end{array}$$

Para se obter a média das duas notas (F e E) fomos obrigados a torná-las equivalentes às do F.N., por meio de regra de três simples:

$$\begin{array}{ccc}
 \text{a) NFN (máxima)} & & F \text{ (máxima)} \\
 10 \xrightarrow{\hspace{10em}} 20 & & \\
 \alpha \xleftarrow{\hspace{10em}} F \text{ (obtida no teste)} & &
 \end{array}$$

$$\begin{array}{ccc}
 \text{b) NFN (máxima)} & & E \text{ (máxima)} \\
 10 \xrightarrow{\hspace{10em}} 1 & & \\
 \beta \xleftarrow{\hspace{10em}} E \text{ (obtida no teste)} & &
 \end{array}$$

Para melhor interpretação, descrevemos abaixo como procedemos com o aluno nº 1.

$$\begin{array}{l}
 \text{Valores obtidos:} \quad (F = 10 \\
 \quad \quad \quad \quad \quad (E = 0,83
 \end{array}$$

$$\begin{array}{ccc}
 \text{a) NFN (máxima)} & & F \text{ (máxima)} \\
 10 \xrightarrow{\hspace{10em}} 20 & & \\
 \alpha \xleftarrow{\hspace{10em}} F = 10 & & \\
 \alpha = 5 & &
 \end{array}$$

b) NFN (máxima)

E (máxima)

10 \longrightarrow 1 $\beta \longleftarrow E = 0,83$

$$\beta = 8,3$$

Os valores obtidos para x e y são, respectivamente, a fluência e exatidão, já transformados para a escala usada pelo falante nativo. Desta maneira, com uma simples média aritmética, obtemos o valor final para comparação:

$$N73 (x) = M = \frac{\alpha + \beta}{2} = \frac{F + E}{2} = \frac{5 + 8,3}{2} = 6,6$$

4.2.2. Notas obtidas com N73

Apresentamos assim o quadro abaixo, que mostra cada aluno com suas respectivas notas (F e E), notas transformadas e a média.

Aluno	Notas	Notas transformadas	Média
1	F = 10,00	5,00	6,60
	E = 0,83	8,30	

(continua)

Aluno	Notas	Notas transformadas	Média
2	F = 13,00	6,50	6,70
	E = 0,69	6,90	
3	F = 7,00	3,50	4,75
	E = 0,60	6,00	
4	F = 9,50	4,75	5,80
	E = 0,69	6,90	
5	F = 12,5	6,25	7,32
	E = 0,84	8,40	
6	F = 9,50	4,75	6,60
	E = 0,86	8,60	
7	F = 14,50	7,25	6,40
	E = 0,57	5,70	
8	F = 6,50	3,25	4,12
	E = 0,50	5,00	
9	F = 9,00	4,50	6,80
	E = 0,91	9,10	
10	F = 10,50	5,25	6,12
	E = 0,70	7,00	
11	F = 9,20	4,60	5,20
	E = 0,58	5,80	
12	F = 12,50	6,25	7,90
	E = 0,96	9,60	

(continua)

Aluno	Notas	Notas transformadas	Média
13	F = 10,50 E = 0,80	5,25 8,00	6,65
14	F = 11,50 E = 0,67	5,75 6,70	6,22
15	F = 8,00 E = 0,67	4,00 6,70	5,32
16	F = 8,50 E = 0,64	4,25 6,47	5,36
17	F = 7,00 E = 0,73	3,50 7,30	5,40
18	F = 7,50 E = 0,65	3,75 6,50	5,12
19	F = 8,00 E = 0,62	4,00 6,20	5,10
20	F = 8,00 E = 0,73	4,00 7,30	5,65
21	F = 6,00 E = 0,72	3,00 7,20	5,10
22	F = 7,00 E = 0,92	3,50 9,20	6,35
23	F = 5,00 E = 0,60	2,50 6,00	4,25

(continua)

Aluno	Notas	Notas transformadas	Média
24	F = 7,00	3,50	4,85
	E = 0,62	6,20	
25	F = 8,00	4,00	6,25
	E = 0,85	8,50	
26	F = 4,00	2,00	4,10
	E = 0,62	6,20	
27	F = 10,00	5,00	4,60
	E = 0,42	4,20	
28	F = 7,00	3,50	3,25
	E = 0,30	3,00	
29	F = 12,00	6,00	6,30
	E = 0,66	6,60	
30	F = 8,50	4,22	4,82
	E = 0,54	5,40	
31	F = 6,00	3,00	4,10
	E = 0,52	5,20	
32	F = 9,00	4,50	5,20
	E = 0,59	5,90	
33	F = 5,50	2,75	5,12
	E = 0,75	7,50	
34	F = 6,50	3,25	5,27
	E = 0,73	7,30	

(continua)

Aluno	Notas	Notas transformadas	Média
35	F = 7,00	3,50	4,75
	E = 0,60	6,00	
36	F = 7,00	3,50	4,60
	E = 0,57	5,70	
37	F = 9,50	4,75	4,90
	E = 0,52	5,20	
38	F = 5,00	2,50	3,75
	E = 0,50	5,00	
39	F = 7,00	3,50	5,30
	E = 0,71	7,10	
40	F = 14,00	7,00	8,20
	E = 0,94	9,40	

4.3. APRESENTAÇÃO DOS DADOS

Após a demonstração dos resultados do N73, apresentamos a comparação com os valores obtidos do FN, confrontando-os no Quadro que segue:

Aluno	N-N73	NFN
1	6,60	6
2	6,70	5
3	4,75	4
4	5,80	5
5	7,32	6
6	6,60	5
7	6,40	6
8	4,12	5
9	6,80	6
10	6,12	5
11	5,20	4
12	7,90	7
13	6,65	6
14	6,22	6
15	5,32	5
16	5,36	5
17	5,40	4
18	5,12	5
19	5,10	4
20	5,65	4
21	5,10	4
22	6,35	5
23	4,25	5
24	4,85	4
25	6,25	5
26	4,10	4
27	4,60	5
28	3,25	4
29	6,30	5
30	4,82	4

(continua)

Aluno	N-N73	NFN
31	4,10	4
32	5,20	4
33	5,12	5
34	5,27	5
35	4,75	6
36	4,60	4
37	4,90	5
38	3,75	5
39	5,30	5
40	8,20	6

4.4. APLICAÇÃO DO COEFICIENTE DE CORRELAÇÃO POR POSTOS DE KENDALL - T (TAU)

Reordenamos os indivíduos de modo que o primeiro conjunto de postos ($N73 = X$) se apresente na ordem natural, isto é, na ordem crescente, como observamos abaixo:

Aluno	N73 = X	NFN = Y
28	3,25	4
38	3,75	5
		{continua}

Aluno	N73 = X	NFN = Y
26	4,10	4
31	4,10	4
8	4,12	5
23	4,25	5
27	4,60	5
36	4,60	4
35	4,75	6
3	4,75	4
30	4,82	4
24	4,85	4
37	4,90	5
19	5,10	4
21	5,10	4
18	5,12	5
33	5,12	5
32	5,20	4
11	5,20	4
34	5,27	5
39	5,30	5
15	5,32	5
16	5,36	5
17	5,40	4
20	5,65	4
4	5,80	5
10	6,12	5
14	6,22	6
25	6,25	5
29	6,30	5
22	6,35	5
7	6,40	6

(continua)

Aluno	N73 = X	NFN = Y
1	6,60	6
6	6,60	5
13	6,65	6
2	6,70	5
9	6,80	6
5	7,37	6
12	7,90	7
40	8,20	6

4.4.1. Cálculo de S

S é o coeficiente que nos dá a intensidade do crescimento da variável $Y = (NFN)$, estando a variável $X = (N73)$ em ordem totalmente crescente.

Cálculo de S

$$\begin{aligned}
 S = & (27 - 0) + (9 - 12) + (26 - 0) + (26 - 0) + (9 - 10) \\
 & + (9 - 10) + (9 - 10) + (23 - 0) + (1 - 23) + (22 - 0) \\
 & + (22 - 0) + (22 - 0) + (8 - 6) + (21 - 0) + (21 - 0) \\
 & + (8 - 4) + (8 - 4) + (19 - 0) + (19 - 0) + (8 - 2) \\
 & + (8 - 2) + (8 - 2) + (8 - 2) + (15 - 0) + (15 - 0) \\
 & + (8 - 0) + (8 - 0) + (1 - 5) + (7 - 0) + (7 - 0) \\
 & + (7 - 0) + (1 - 2) + (1 - 2) + (5 - 0) + (1 - 1) \\
 & + (4 - 0) + (1 - 0) + (1 - 0) + (0 - 1)
 \end{aligned}$$

$$\begin{aligned}
 S = & 27 + (-3) + 26 + 26 + (-1) + (-1) + (-1) + 23 \\
 & + (-29) + 22 + 22 + 22 + 2 + 21 + 21 + 4 + 4 + 19 \\
 & + 19 + 6 + 6 + 6 + 6 + 15 + 15 + 8 + 8 + (-4) + 7 \\
 & + 7 + 7 + (-1) + (-1) + 5 + 0 + 4 + 1 + 1 + 1 = 361 \\
 & + (-41) = 325
 \end{aligned}$$

$$S = 325$$

Encontrado o S aplicaríamos a seguinte fórmula, caso não houvesse empates nos escores:

$$T = \frac{S}{1/2 N (N - 1)}$$

Como neste caso houve vários empates, atribuímos às observações empatadas a média dos postos que lhes caberiam se não houvesse empates. O efeito do empate consiste em modificar o denominador da fórmula de T da seguinte maneira:

$$T = \frac{S}{\frac{\sqrt{1/2 N(N-1)} - T_X}{\sqrt{1/2 N(N-1)} - T_Y}}$$

onde:

$$T_X = 1/2 \sum t (t-1) \text{ e } T_Y = 1/2 \sum t (t-1),$$

sendo t o número de observações empatadas em cada gru-

po de empates.

4.4.2. Cálculo de T_x e T_y

Na variável x , houve empates nos postos 4,10, 4,60, 4,75, 5,10, 5,12, 5,20 e 6,60. Em cada um destes casos o número de empates foi dois, ou seja, sete conjuntos de postos empatados.

$$T_x = 1/2 \sum T (T - 1)$$

$$T_x = 1/2 \left[2(2-1) + 2(2-1) + 2(2-1) + 2(2-1) + 2(2-1) + 2(2-1) + 2(2-1) \right]$$

$$T_x = 1/2 \cdot 14$$

$$T_x = 7$$

Na variável y , houve empates nos postos 4, 5 e 6, da seguinte maneira:

No posto 4 o nº de empates foi $t = 13$

No posto 5 o nº de empates foi $t = 18$

No posto 6 o nº de empates foi $t = 8$

$$T_y = 1/2 \sum T (T - 1)$$

$$T_y = 1/2 \left| 13 (13-1) + 18(18-1) + 8(8-1) \right|$$

$$T_y = 1/2 \cdot 518$$

$$T_y = 259$$

4.4.3. Cálculo de T e Prova de Significância (Z)

Tendo-se obtido os valores de S, T_x e T_y e sabendo-se que N = 40, substituímo-los na fórmula abaixo:

$$T = \frac{S}{\sqrt{1/2 N(N-1)} - T_x \cdot \sqrt{1/2 N(N-1)} - T_y}$$

$$T = \frac{325}{\sqrt{1/2 \cdot 40(40-1)} - 7 \cdot \sqrt{1/2 \cdot 40(40-1)} - 259}$$

$$T = 0,512$$

Este valor representa o grau de correlação ou relacionamento entre os 40 pares de notas obtidas pelo teste aplicado (N73) e pela apreciação do professor (NFN).

Podemos assim verificar qual o nível de significância (Z) deste resultado, utilizando-nos da seguinte fórmula:

$$Z = \frac{T}{\frac{2(2N + S)}{9N(N-1)}}$$

$$Z = \frac{0,512}{\sqrt{\frac{2(2,40 + 5)}{9,40 (40-1)}}$$

$$Z = 4,65$$

Referindo-nos à Tábua A, vemos que $Z \geq 4,65$ tem probabilidade $p = 0,00003$ de ocorrência sob H_0 . Podemos assim rejeitar H_0 ao nível de significância $\alpha = 0,00003$ e concluir que existe correlação entre as duas variáveis.

TABLE A. Tabla de probabilidades asociadas con valores tan extremos como los valores observados de z en la distribución normal

El cuerpo de la tabla contiene las probabilidades de una cola conforme a H_0 de z . La columna marginal izquierda tiene diferentes valores de z hasta de una cifra decimal. La hilera superior tiene diferentes valores hasta la segunda cifra decimal. Así por ejemplo, la p de una cola de $z \geq 0.11$ o $z \leq -0.11$ es $p = 0.4562$.

[illegible]

5. CONCLUSÃO

Com a realização deste trabalho pretendíamos verificar se o teste N73 é tão confiável e válido para medir a comunicação oral quanto a entrevista realizada por falantes nativos. Para tanto, testamos quarenta alunos pelos dois métodos e fizemos uma comparação entre os resultados, usando para isso um método de comparação estatística.

Antes, porém, para um melhor embasamento teórico, fizemos um levantamento sobre testes. Em primeiro lugar, verificamos os diversos conceitos defendidos pelos lingüistas, a fim de que encaixássemos o presente teste dentro dos mais adequados. Finalizada esta parte geral, partimos então para um relato sobre os diversos testes usados para medir a comunicação oral.

A partir deste embasamento teórico, sentimos a dificuldade existente que gira em torno de como e qual a melhor maneira de testar a comunicação oral. Pudemos verificar que esta dificuldade deve-se aos fatos de ser um campo em que as pesquisas são em pequeno núme-

ro e à própria dificuldade que o assunto apresenta. Vimos, então, que não se chegou a uma conclusão unânime a respeito da definição do termo proficiência; sendo que ao testar a proficiência oral, cada examinador ou escola toma como base a sua própria definição de proficiência.

É necessário esclarecermos que ao aplicarmos e analisarmos o teste *N73 Instantaneous Evaluation of Speaking Ability* não almejávamos apresentá-lo como o teste ideal para medir a proficiência oral, mas somente verificar se poderia ser utilizado da mesma forma que a entrevista oral.

Desta maneira, após a aplicação dos dois testes, realizou-se uma comparação, utilizando o coeficiente de correlação por postos de Kendall. O resultado obtido por este método nos diz se existe ou não correlação entre duas variáveis. No presente experimento, estas variáveis são os resultados obtidos com a aplicação de ambos os testes: N73 e falante nativo. O resultado alcançado, após a utilização do método, fizemos concluir que existe correlação entre as duas variáveis. A existência de correlação nos conduz à comprovação da hipótese de que o teste N73 é tão adequado quanto o teste realizado pelo falante nativo, para medir a habilidade oral da fala.

6. REFERÊNCIAS BIBLIOGRÁFICAS

1. ALLEN, J.P.B. & DAVIES, A. Testing and Experimental Methods. London, Oxford University Press, 1977. 4 v.
2. BEARDSMORE, H. Baetens. Testing Oral Fluency. IRAL. Alemanha, 12(4):317-25, nov.1974.
3. BEARDSMORE, H. Baetens & RENKIN, A. A Test of Spoken English. IRAL. Alemanha, 19(1):1-11, fev. 1971.
4. BANATHY, Bela & LANGE, Dale L. A Design for Foreign Language Curriculum. Lexington, D.C. Heath, 1972. 128 p.
5. CLARK, John L.D. Theoretical and Technical Considerations in Oral Proficiency Testing. In: JONES, R.L. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.10-28.
6. DAVIES, Alan. Language Testing. Language Teaching & Linguistics: Abstracts. London, 11(3):145-59, jul.1978.
7. _____. Language Testing - Part II. Language Teaching & Linguistics: Abstracts. London, 11(4):215-26, oct.1978.
8. FERGUSON, Nicolas. Test N73: Instantaneous Evaluation of Speaking Ability. IRAL, Alemanha, 26(4):340-49, nov.1978.
9. GROOT, Peter J.M. Testing Communicative in Listening Comprehension. In: JONES, R.L. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.45-58.

10. HERRIOT, Peter. An Introduction to the Psychology of Language. 2.ed. London, Methuen & Co. Ltd, 1971. 189 p.
11. JAKOBOVITS, Leon A. A Psycholinguistic Analysis of the Issues. 2.ed. Mass., Newbury House, 1971. 336 p.
12. JONES, R.L. Testing Language Proficiency in the United States Government. In: JONES, R. L. & SPOLSKY, B. Testing Language Proficiency. U.S. A., Center for Applied Linguistics, 1975. p.1-9.
13. LADO, Robert. Language Testing. In: Language Teaching. U.S.A., McGraw Hill, 1964. p.158-70.
14. LEESON, Richard. Fluency and Language Teaching. London, Longman, 1975. 202 p.
15. MACKEY, W.F. The Measurement of Language Learning. In: _____. Language Teaching Analysis. London, Longman, 1966. p.403-16.
16. OLLER Jr., John W. Discrete Point, Integrative, or Pragmatic Tests. In: _____. Language Tests at School. London, Longman, 1979. p.36-69.
17. _____. Tests of Productive Oral Communication. In: OLLER Jr., John W. Language Tests at School. London, Longman, 1979. p.303-35.
18. _____. The Ultimate Problem of Test Validity. In: OLLER Jr., John W. Language Tests at School. London, Longman, 1979. p.403-08.
19. ROBINSON, Peter. Oral Expression Tests: 1. English Language Teaching. London, 25(2):151-55, jan. 1971.
20. _____. Oral Expression Tests: 2. English Language Teaching. London, 25(3):260-66, jun.1971.
21. SEWARD, B.H. Measuring Oral Production in E.F.L. English Language Teaching Journal. London, 28(1):76-80, nov.1973.
22. SIEGEL, Sidney. O Coeficiente de Correlação por Postos de Kendall: T. In: _____. Estatística não-paramétrica. S.P., McGraw-Hill, 1979. p.241-51.
23. TOWNSON, M. Testing Oral Skills at University Level. English Language Teaching. London, 27(2): 199-

205, feb.1973.

24. TTOFI, C. & CREED, T.S. English in Mind. London, Macmillan, 1974. 194 p.
25. WILDS, Claudia P. The Oral Interview Test. In: JONES, R.L. & SPOLSKY, B. Testing Language Proficiency. U.S.A., Center for Applied Linguistics, 1975. p.29-44.
26. ZIERER, E. The Test of Aural Perception in Foreign Language Teaching from the Standpoint of Information Theory. IRAL, Alemanha, 19(2):125-30, maio 1971.